(51) International Patent Classification 4 : C12P 21/00, C12N 15/00, 9/58 C12N 5/00, 1/20, C07H 15/12 C11D 7/42	**A1**	(11) International Publication Number: **WO 88/07581** (43) International Publication Date: 6 October 1988 (06.10.88)

(21) International Application Number: PCT/US88/01040

(22) International Filing Date: 28 March 1988 (28.03.88)

(31) Priority Application Number: 035,816

(32) Priority Date: 3 April 1987 (03.04.87)

(33) Priority Country: US

(71) Applicant: AMGEN INC. [US/US]; 1900 Oak Terrace Lane, Thousand Oaks, CA 91320 (US).

(72) Inventors: SAMAL, Babru, B. ; 4017 San Luis Street, Newbury Park, CA 91320 (US). STABINSKY, Yitzhak ; 3415 Heidelberg Drive, Boulder, CO 80303 (US).

(74) Agents: SIMONTON, Pamela, A. et al.; Amgen Inc., 1900 Oak Terrace Lane, Thousand Oaks, CA 91320 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: NOVEL PROTEOLYTIC ENZYMES

(57) Abstract

This disclosure relates to a novel class of serine proteases isolated from a culture medium of fungus *Tritirachium album.* The serine proteases disclosed have a high degree of stability in detergent formulations. In addition, this disclosure relates to a process for producing such serine proteases using recombinant techniques.

- 1 -

NOVEL PROTEOLYTIC ENZYMES

BACKGROUND

5   This invention relates to novel serine proteases
     isolated from a culture medium of the fungus
     Tritirachium album. The serine proteases of the present
     invention have a high degree of stability in aqueous
     solutions and dry detergent formulations. The invention
10   further relates to detergent compositions containing
     such proteases and to the use of the proteases in
     detergents and cleaners or spot cleaners. In addition,
     the present invention further relates to DNA sequences
     encoding for the proteases and to a method for producing
15   the proteases.

     Serine proteases are proteolytic enzymes having a serine
     residue at their active site. Modification of the
     active site serine residue by agents such as phenylmethyl-
20   sulfonylfluoride (PMSF) inactivates these enzymes. There
     are two classes of serine proteases, chymotrypsin-like
     proteases and subtilisin-like proteases. Subtilisins are
     serine proteases which generally act to cleave internal
     peptide bonds of proteins or peptides. Subtilisins are
25   secreted by a number of Bacillus species and extensively
     used commercially (see U.S. patent No. 3,623,957, J.
     Miller, 1970, J. Appl. Bacteriol., 33, 207; Ward, O.P.
     1983, pp. 251-317, Enzymes and Biotechnology, ed., W.M.
     Fogarty, Applied Science Publishers, London).
30   Subtilisins have been utilized in a number of detergent
     formulations (see U.S. Patent Nos. 1,240,058, 3,749,671;
     3,790,482; 4,266,031, U.K. Patent No. 1315937).

     The use of proteases in industrial processes which
35   require hydrolysis of protein has been limited due to
     the enzyme instability under the temperature and pH

- 2 -

conditions associated with such processes. Although
thermal inactivation of the protease may be the most
important factor in restricting the industrial use of a
protease, other factors such as lack of effectiveness
5   over broad pH ranges and use of denaturing agents in
detergent formulations may also have a detrimental
effect regarding the use of proteases in industrial
processes. The known Bacillus derived subtilisins are
not ideal for all applications as detergent enzymes, in
10   particular, application requiring greater storage
stability and activity at broader ranges of pH and
temperature. Therefore there is a need for a class of
proteases that are characterized by high stability with
respect to temperature, pH, denaturing agents and the
15   like.

There also is a need for proteases that are compatible
with detergents and have sufficient shelf-life in liquid
detergent formulations to be commercially practical.
20   Thermostable fungal serine proteases have been evaluated
in detergent applications, including proteases obtained
from Tritirachium album (Ebeling, W. et al., 1971, German
Offenbach, 1965, 281), Malbranchea pulchella (Ong, P.S. et
al., Can. J. Microbiol. 22, 165), Acremonium kiliense,
25   Fusarium and Gibberella spp (Isono, M., et al., 1972,
U.S. Patent No. 3652399). Proteinase K (EC 3, 4, 21, 14)
was isolated from Tritirachium album (Ebeling, W. et al.,
1974, Eur. J. Biochem. 47, 91-97) and has been extensively
studied by different groups (Kraus, E. et al., 1976,
30   Hoppe Seyler's Z. Physiol. Chem. 357, 937-947, ibid. 357,
233-237, and Morihara, K. et al., 1975, Agr. Biol. Chem.
39, 1489-1492). The three dimensional structure of
proteinase K is similar to that of subtilisins (Paehler,
A. et al., 1983, EMBO J. 3, 1311-1314) and there is about
35   35% homology of the amino acid sequence of proteinase K
with that of subtilisins. (Jany, K-D., et al., FEBS

- 3 -

Letters, 199, 139-144). Detergent compatibility of
proteinase K has been suggested from its activity in the
presence of high concentration of detergents (Hilz, H.
et al., 1975, Eur. J., Biochem., 56, 103-108).

5   Proteolytic enzymes generally catalyze the cleavage of
peptide bonds only within a certain range of pH and
temperature. Moreover, even under optimal conditions, a
proteolytic enzyme retains its activity only when its
10   polypeptide chain is in native conformation. Unfolding
of the native structure often occurs when the enzyme is
exposed to extremes of pH or temperature or to certain
detergent additives such as surfactants and metal-
chelating agents. ·The latter exert their effect
15   especially on enzymes that require metal ions such as
Ca^{2+} for stabilizing their native structure, i.e.,
bacterial subtilisin. For proteases, partial unfolding
of the native conformation of the enzyme may lead to
acceleration of autodigestion and therefore, to
20   irreversible enzyme inactiviation. Because most commer-
cial laundry detergents have an alkaline pH, it is
desirable that the enzyme utilized in such detergents be
active and stable in a pH range of between 7.5 and 13
and in a temperature range of between 20-65°C. Moreover,
25   it is desirable that the activity of such enzymes be
relatively independent of calcium and magnesium ions and
be compatible with surfactants and sequestrant builders.
Bacterial serine proteases of the subtilisin family
fulfill these requirements to some extent, however their
30   stability in liquid detergent formulation is limited.

                        SUMMARY OF THE INVENTION

The present invention relates to novel serine proteases
35   that are characterized by improved stability in
detergent formulations. In particular, the present

- 4 -

invention describes the isolation, purification and
characterization of novel serine proteases obtained from
the strain of fungus Tritirachium album Limber (ATCC
22563). The serine proteases of the present invention
5   have been found to be superior detergent enzymes and
possess a high degree of thermal stability in aqueous
solutions particularly at elevated temperatures, thereby
making the proteases suitable for use as additives in
commercial liquid detergent formulations. The present
10   invention further relates to the isolation and
characterization of the genes encoding such proteases.
The invention further relates to the use of the serine
proteases of the present invention in detergents and
cleaners or spot cleaners and compositions containing
15   the serine proteases.

In addition, the present invention relates to the use of
an oligonucleotide probe which hybridizes with
complementary DNA sequences in the genomic or cDNA
20   clones of the serine proteases disclosed herein.
Finally the present invention relates to DNA sequences
useful in securing expression in a procaryotic or
eucaryotic host cell a serine protease isolated from
culture media containing strain of T. album Limber (ATCC
25   22563) and the method for isolating such proteases.

The present invention further relates to a purified and
isolated serine protease having the structural
conformation (i.e., continuous sequence of amino acid
30   residues) of a serine protease isolated from a culture
of Tritirachium album Limber strain ATCC 22563 and
characterized by being the product of procaryotic or
eucaryotic expression of an exogenous DNA sequence.

35   Also, the present invention provides a process for the
production of a serine protease having the structural

- 5 -

conformation of a serine protease isolated from a
culture or <u>Tritirachium</u> <u>album</u> Limber strain ATCC 22563,
said process comprising growing under suitable nutrient
condition procaryotic or eucaryotic host cells
5   transformed or transfected with a DNA vector including a
DNA sequence useful in securing expression in a
procaryotic or eucaryotic host cell the desired serine
protease and isolating desired serine protease of the
expression of DNA sequences in said vector.
10
Also provided herein are oligonucleotide probes capable
of hybridizing with a DNA sequence capable of expressing
in a procaryotic or eucaryotic host cell a serine
protease having the structural conformation of a serine
15   protease isolated from culture of media of <u>Tritirachium</u>
<u>album</u> Limber strain ATCC 22563.

BRIEF DESCRIPTION OF THE DRAWINGS

20   Figure 1 represents the nucleotide sequence of the
genomic clone for protease TW7;

Figure 2 represents the nucleotide sequence of the cDNA
clone for protease TW7;
25
Figure 3 represents the nucleotide sequence of the
coding strand, correlated with the amino acid sequence
of the protease TW7;

30   Figure 4 represents a comparison of the amino acid
sequences of protease TW7 with those of proteinase K,
subtilisin novo, subtilisin Carlsberg, subtilisin DY and
thermitase;

35   Figure 5 represents the pH profiles of protease TW7;

- 6 -

Figure 6 represents the temperature profiles of the
protease TW7;

Figure 7 represents the nucleotide sequence of the cDNA
5   clone for protease TW3;

Figure 8 represents the nucleotide sequence of the
coding strand, correlated with the amino acid sequence
of the protease TW3;

10

Figure 9 represents a comparison of the amino acid
sequences of protease TW3 with those of proteinase K,
subtilisin novo, subtilisin Carlsberg, subtilisin DY and
thermitase;

15

Figure 10 represents the pH profiles of protease TW3;

Figure 11 represents the temperature profiles of the
protease TW3; and

20

Figure 12 represents a flow diagram illustrating the
components utilized in construction of pCFM 1156 TW7.

- DETAILED DESCRIPTION

25

Because T. album Limber is a slow growing fungus capable
of producing only low level quantities of proteases,
production of proteinase K or any other subtilisin-like
enzymes in large commercial quantities by fermenting the
30   fungus, is not practical. In addition, the growth of
the fungus T. album as well as the production of
proteases are slow, making the commercialization of a
subtilisin-like enzyme uneconomical.

35   One aspect of the present invention relates to the
isolation of a gene encoding the novel serine proteases

- 7 -

of the present invention from T. album Limber, and the
cloning and expression of such a gene in a suitable
microorganism. The gene was isolated using a
deoxyribonucleotide oligomer that hybridizes with a gene
5 having a DNA sequence encoding the amino acid residues
around the active serine residue. Using this approach,
more than one gene has been isolated from the genomic
library. This suggests the production of more than one
serine protease from T. album Limber. The amino acid
10 sequences of the amino termini of the proteases were
found to be different. The novel proteases, protease TW7
and protease TW3 were isolated and characterized.

The gene encoding for protease TW7 was isolated from a
15 genomic library. In particular, the isolation of
subtilisin-like genes from the genomic library of T.
album comprises (1) construction of a library from the
genomic DNA of T. album Limber in a pBR322 plasmid; (2)
screening of the genomic library with a labeled
20 oligonucleotide probe which hybridizes specifically with
subtilisin-like genes; (3) restriction enzyme analysis
of positive clones followed by Southern blot hybridi-
zation with various restriction fragments for the
identification of clones carrying the entire gene; (4)
25 subcloning of restrictive fragments carrying the entire
gene in bacteriophage M13 for DNA sequence analysis; (5)
designing oligonucleotide primers based on partial DNA
sequence data to complete DNA sequencing of both strands
of the gene.
30
Sequencing of both strands of the genomic gene revealed
the presence of two introns. In order to express genes
containing introns in microorganisms that lack splicing
enzymes, e.g., B. subtilis, E. coli, etc., it is
35 necessary to reconstruct the gene (i.e., utilizing in
vitro splicing) or obtain the gene from a cDNA

- 8 -

library. The genes encoding for protease TW7 and
protease TW3 were isolated from a cDNA library. The
isolation of subtilisin genes from the cDNA library of
T. album Limber comprises (1) isolating of total RNA
5   from T. album Limber; (2) fractioning of RNA on an
oligonucleotide dT cellulose column and isolation of
polyadenylated mRNA fraction; (3) using an oligonucleotide
probe for Northern blot analysis to confirm the presence
of subtilisin-like mRNA; (4) cDNA synthesis and construc-
10   tion of a cDNA library in a pBR322 derived plasmid; (5)
screening of the cDNA library with a ^{32}P-labeled
oligonucleotide probe that was utilized in Step 3; (6)
isolation and restriction analysis of positive cDNA
clones; (7) subcloning of restriction fragments from
15   positive cDNA clones that carry the entire protein
coding sequence in bacteriophage M13 for DNA sequencing;
(8) using oligonucleotide primers to complete the DNA
sequencing of both strands of the gene.

20   In accordance with the above procedures, two cDNA genes
have been identified whose amino acid sequences as
deduced from the DNA sequences indicated that their
products are subtilisin-like enzymes. One of the gene
products has the amino acid sequence represented in
25   Figure 3 and was named protease TW7. Protease TW7
exhibits approximately 53% amino acid sequence homology
with proteinase K from T. album Limber. The putative
amino acid sequence of protease TW7 indicated that unlike
proteinase K, this protease has a net negative charge and
30   therefore may be separated from the other proteases
using DEAE sepharose or DEAE-cellulose chromatography.
The second of putative gene products was named protease
TW3. The degree of amino acid sequence homology of this
protease and proteinase K from T. album Limber was
35   approximately 90%. The putative amino acid sequence of
protease TW3 indicated that the enzyme has a net

- 9 -

positive charge and therefore can be separated from the
other proteases by CM-cellulose chromatography.

The serine proteases of the present invention were
5   isolated from the culture broth of T. album Limber
strain (CBS 348.55) ATCC Deposit No. 22563 as follows:

The particular fungus T. album strain (ATCC 22563) was
grown in media containing only proteins as nitrogen
10   sources, e.g., skim milk, bovine serum albumin or soy
flour. The culture media were tested for proteolytic
activity using azocasein as the substrate. When the
proteolytic activity reached a plateau or started to
decline, the culture broth was separated from fungal
15   mycelia by centrifugation. Proteins in the supernatant
were precipitated using ammonium sulfate. The
precipitate was dissolved in 20 mM sodium phosphate '
buffer at pH 6.0 and the solution was dialyzed against
the same buffer. The solution was passed through a
20   column of CM-52 (Whatman) to capture proteinase K-like
enzymes and then passed through a DE-52 (Whatman)
column. Proteinase TW7 was eluted from the DE-52 column
with 100 mM NaCl in 20 mM sodium phosphate at pH 6.0.
It was dialyzed against water and then lyophilized.
25   Protease TW3 was bound to CM-52 from which it was eluted
with 200 mM NaCl in 20 mM sodium phosphate at pH 6.0.

The detergent compatibility properties of the serine
proteases of the present invention have been character-
30   ized. As shown in Figures 5 and 10, protease TW7 and
protease TW3 are active over a wide pH range. Although
the maximum activity occurs approximately at pH 10,
protease TW7 has considerable activity within pH range 9
to 13.   In addition, protease TW7 retains at least 45%
35   of its maximum activity present in a buffer having pH of
6.0. As represented in Figure 6, protease TW7 has

- 10 -

enzymatic activity over a broad temperature range. The
maximum activity is between 60-65°C, although at 20°C
protease TW7 retains at least 25% of its original
activity. Protease TW7 is stable at 52°C in buffers of
5   various pH values. Following incubation for a period of
one hour at pH 8 and 10.3, approximately 100% of the
original activity is retained, while under similar
conditions a commercial subtilisin retained only up to
30% of its original activity. In the presence of 0.5%
10   SDS at pH 8.0, protease TW7 retains 100% of its activity
after one hour incubation at 52°C while the commercial
subtilisin retains only about 5-6% of its original
activity.

15   As illustrated in Figure 10, protease TW3 is also active
over a broad pH range. For example, protease TW3 has
its highest enzymatic activity at pH 9, although there
is considerable activity within a pH range of from 7 to
10. As represented in Figure 11, the temperature range
20   of enzymatic activity of protease TW3 is also broad.
The maximum activity is between 55-65°C, although at
25°C, protease TW3 retains at least 32% of its
activity. This enables protease TW3 to be active at a
wide range of washing temperature. In addition,
25   protease TW3 is very stable at 50°C in buffers of
various pH values. Following one hour incubation at pH
4.0, 8 and 10.0, 50-90% of the original activity of
protease TW3 is retained, while under similar conditions
a commercial subtilisin retains only minimal activity.
30
The stability of the serine proteases of the present
invention have been evaluated in different laundry
detergent formulations. Any endogenous enzyme that was
present was inactivated prior to evaluation by heating
35   the detergent formulation at 65°C for one hour.

- 11 -

In addition to the enzyme the commercial washing powder composition of the present invention will generally contain:

(a) At least one surfactant which may be anionic, non-ionic, or amphoteric, or a water-soluble soap. Typically, an anionic surfactant (e.g. a linear alkyl aryl sulphonate) is used in admixture with a non-ionic (e.g. an alkyl phenyl polyglycol ether) in amounts of 5-30 and 1-5 percent by weight, respectively, of the washing composition.

(b) One or more builders, preferably having a concomitant sequestering function. Sodium tripolyphosphate, sodium citrate, sodium silicate, and zeolites are examples of such compounds, usually constituting from 10 to 70 percent by weight of the detergent composition.

(c) A bleaching agent, preferably a peroxy compound such as sodium perborate, typically incorporated in an amount up to 30 percent by weight of the composition.

(d) Ancillary agents, such as carboxymethyl cellulose, optical brighteners and perfumes. If required, a pH-adjusting agent is added to give a pH of the laundering medium in the range of from 8.0 to 10.5.

The particulate protease preparation of the invention is added in an amount calculated to give a protease activity of at least 0.1 Anson units (AU, vide infra), preferably 0.5-2.5 AU per 100 g of washing composition. If required, balance to 100 percent may be established with an inorganic filler, preferably sodium sulphate.

Liquid detergent compositions may be prepared from enzyme slurries, preferably in non-aqueous media. Typically, such slurries may consist of a suspension of finely ground protease concentrate in a liquid non-ionic surfactant, for example Tergitol 15 S 9 or a mixture of

- 12 -

such surfactants. Usually, the slurry will also contain
one or more inorganic fillers, such as finely ground
sodium chloride, optionally in admixture with a
suspension stabilizer, for example fumed silica (Aerosil
5   200). Tergitol and Aerosil are trade marks.

The protease slurry of the invention is added in an
amount calculated to give a protease activity of at
least 0.1 AU preferably 0.5-2.5 AU per 100 g of liquid
10   detergent composition.

The washing compositions may be prepared in the usual
manner, for example by mixing together the components.
Alternatively, a pre-mix is made, which is then mixed
15   with the remaining ingredients.

The following Examples will further serve to illustrate
the invention although it will be understood that the
invention is not limited to these specific examples.
20

Because of the good stability and activity properties
described, the proteolytic enzyme according to the
invention can be used in all fields where proteolytic
enzymes are generally used. In particular, it can be
25   used for detergents and cleansers or spot removers, as a
depilatory in tanning, and also in the food industry for
the preparation of protein hydrolysates and in serology
for the detection of incomplete antibodies. It is
particularly advantageous for use in the food industry
30   and in serology that the enzyme according to the
invention has such an excellent stability in the solid
or dissolved form that physiologically acceptable
quantities of calcium ions may not be necessary to
stabilize the enzyme in aqueous solutions, in contrast
35   to those of other enzyme preparations.

- 13 -

Example 1

Production of serine protease from Tritirachium album
Limber.

The fungus, Tritirachium album Limber (CBS348.55, ATCC
22563) was obtained from American Type Culture
Collection, 912301 Parklawn Dr., Rockville, MD. The
fungus belongs to the family Moniliaceae (Limber, 1940,
Mycologia, 32,23-30). Typical features of the fungus
include hyaline mycelium, sparingly branched hyaline
conidia, pure white dense mycelliar mat forming low dome
or hemisphere. According to the Centralbureau voor
Schimmelcultures (Oosterstraat 1, Baarn, P.O. Box 273,
3740 Ag Baarn, Netherland, ref. MAAS/tvs/714, dated
October 4, 1985) the strain was originally isolated from
deceased skin. This strain is not the same strain (CBS
747.69) from which proteinase K was isolated.
Tritirachium album was propagated on malt-peptone-agar
plates that contained 3% malt extract, 0.3% casein
peptone and 2% bactoagar. The fungus was allowed to
sporulate at room temperature for 7-8 days. Spores were
collected in 15-20 ml of sterile distilled water and
were maintained at room temperature for at least 30 min.
before inoculation into a liquid media.

The composition of the liquid media for the production
of proteases was either 2% skim milk and 0.17% yeast
nitrogen base (Cat. No. 0335-15, obtained from Difco
Laboratories, Detroit, MI) or 1% bovine serum albumin
(BSA), 1% glucose and 0.17% yeast nitrogen base. The
BSA containing mediam was sterilized by passing through
a 0.45µ filter while the skim milk containing media was
autoclaved for 30 minutes.

- 14 -

500 ml of each media was inoculated with the spore
suspension and the resulting mixture was shaken. Four
to five drops of antifoam were used to suppress foaming
during shaking. The production of extramycelial

5   proteases was monitored using either N-succinyl-alanyl-
alanyl-prolyl-phenylalanine nitroanilide (Delmar et al.,
Anal Biochem. 99, 316) or azocasein (Charney, J. et al.
1947. J. Biol. Chem. 177, 501) as substrate. It was
found that the production of protease was media-

10   dependent. In skim milk media, the largest amount of
protease was produced within 8-9 days of culture, while
in BSA media the maximum production of protease did not
occur until after 14-15 days of culture. In both media,
protease activity declined after achieving maximum

15   production level.

Example 2

Design and synthesis of a probe for detection of
20   subtilisin-like genes.

Subtilisin-like enzymes share a high degree of sequence
homology around their active site serine residue.
Serine at position 221 has previously been identified as

25   the reactive serine in serine proteases. Moreover, the
assignment of a serine protease to the subtilisin family
is based on the amino acid sequence
                    GLY-THR-SER-MET-ALA
Amino acid sequence alignment of several subtilisin-like

30   enzymes revealed that the homology stretches beyond this
sequence as follows:

GLY-THR-SER-MET-ALA-$^{SER}_{THR}$-PHE-HIS-VAL-ALA-GLY-$^{LEU}_{ALA}$-ALA-ALA

35   (Stahl, M.L. et al., 1984, supra; Wells, J.A. et al.,
1983, supra; Vasantha, N. et al., 1984, supra; Svendsen,

- 15 -

I. et al., 1986, FEBS Letters 196, 228-232; Koide, Y. et
al., 1986, J. Bacteriol. 167, 110-116; Kaneda, M. et
al., 1984; J. Biochem. 95, 825-825; Jany, K.-D. et al.,
1985, Biol. Chem. Hoppe Seyler 366, 485-492).
5   Analysis of the DNA sequences encoding the amino acids
in subtilisins also showed high degree of
conservation. Based on these observations, the
following deoxyoligonucleotide (41-mer) was designed for
probing genes that contain similar coding sequences:
10

5' GCT GCT AIT CCG GCA ACG TGA GGA GTC GCC ATG GAC GTT CC 3'

The probe was synthesized using the phosphotriester
method of Beaucage et al. (1981, Tetrahedron Letters 22,
15   1859-1862).

This oligonucleotide probe hybridizes with complementary
DNA sequences in the genomic clone or cDNA clone of the
serine protease at a temperature range of from 50-78°C
20   depending on the number of mismatches. Therefore, when
using this probe, it is necessary to include hybridiza-
tion and washing at several temperatures and salt
concentrations for adequate stringency. The 41-mer
oligonucleotide may be used for probing genes encoding
25   for subtilisin-like enzymes in genomic and cDNA libraries
as well as mRNA encoding subtilisin-like enzymes.

Example 3

30   Construction of the genomic library and isolation of the
gene for protease TW7.

Mycelia were collected on miracloth, as previously
described, following 9-15 days of fungal growth.
35   Mycelia were weighed and quickly frozen in dry ice and
isopropanol. To 10 g of frozen mycelia was added 10 g

- 16 -

of autoclaved alumina and 20 ml of lysis buffer (4% SDS,
20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 0.15M NaCl). The
resulting mixture was ground for 5 minutes in a sterile
mortar using a pestle. An additional 20 ml of buffer
5   and 40 ml of phenol-chloroform-isoamyl alcohol (25:24:1)
was added to the ground mixture and the lysate was
shaken for 20-30 min at room temperature and centrifuged
for 10 min. The aqueous phase was extracted with phenol
and chloroform. Ribonuclease A and proteinase K were
10   added to the lysate to eliminate RNA and protein
respectively from the DNA preparation. Ethanol was
added to the lysate and DNA was spooled, suspended in TE
(20 mM Tris-HCl, pH 8.0, 1 mM EDTA) and stored at 4°C.

15   A complete digestion of the genomic DNA isolated from T.
album was conducted using restriction enzymes EcoRI and
BamHI to construct a library in plasmid vector pBR322.
The vector pBR322 was also digested with the same
enzymes and then treated with alkaline phosphatase.
20   Following ligation of the T. album DNA fragments with
the pBR322 vector, transformation of competent HB101
cells was conducted in accordance with the procedures of
Mandel et al., 1970, J. Mol. Biol. 53, 159-162.
Ampicillin-resistant colonies were probed for the
25   presence of subtilisin-like genes using the 41-mer
oligonucleotide probe of Example 2. Colony lift and
hybridization was conducted according to the procedures
of Grunstein et al., 1975, Proc. Natl. Acad. Sci. USA,
72, 3961-3965 except that a gene screen membrane (NEF-
30   972, New England Nuclear) was used instead of
nitrocellulose. Filters were processed for DNA
denaturation in situ, followed by neutralization, drying
and baking at 80°C in vacuum. Prehybridization was
conducted at 55°C for 3-4 hours in 5X Denhardt's
35   solution containing 200 µg tRNA/ml. Hybridization was
carried out at 55°C for 20 hours in 5X SSC, 1% SDS, 1X

- 17 -

Denhardt's solution and 200 μg tRNA/ml. The filters
were washed in stringent conditions, 2X SSC at 55°C
until background radioactivity was negligable. A second
round of screening of putative positive colonies with
5   the 41-mer oligonucleotide probe was conducted by
reinoculating the putative positive clones onto L agar
plates containing 50 μg ampicillin/ml and repeating the
screening steps described above. Following the second
round of screening, only one positive clone was
10   obtained. The positive colony was cultured in Luria
broth containing 50 μg ampicillin/ml. From the
overnight culture, a plasmid containing the gene for
protease TW7 was isolated using the procedures described
by Birnboim, Methods in Enzymol. 100, 243-154, (1983).
15   For purposes of Southern blotting and extensive
restriction mapping, the plasmid was purified using
cesium chloride ethidium bromide gradients and
ultracentrifugation. A 2.8 kb fragment of T. album DNA
was found to contain the gene for protease TW7.

20

Restriction fragments of this plasmid generated by
different enzymes were resolved on agarose gels by
electrophoresis, transferred onto gene screen plus®
membranes then probed with the 41-mer oligonucleotide
25   probe of Example 2 to identify the fragment containing
the complete gene for protease TW7. From this, it was
determined that an EcoRI-ClaI fragment containing 1056
nucleotides contained the gene. This fragment was then
subcloned in the bacteriophage M13mp18 and M13mp19 for
30   single strand DNA sequencing using the universal primer
(Messings, J. 1983; Methods in Enzymol. 101C, 20-75).
The dideoxy chain termination procedure described by
Sanger, F. et al., (1977), Proc. National Acad. Sci.
USA, 74, 5463 was used for DNA sequencing. Once some
35   partial DNA sequence was obtained, additional
oligonucleotide primers were used to complete the

- 18 -

sequencing of the two strands of the gene. The
nucleotide sequence of the gene thus characterized is
represented in Figure 1. The gene contains 1056
nucleotides and is defined by a site for restriction
5   enzyme EcoRI on the 5' end and a ClaI on the 3' end.
Restriction mapping of the fragment reveals the presence
of unique sites for the restriction enzymes HindIII,
KpnI and BglI. The putative protease encoded by this
gene was named protease TW7.
10

The isolated gene for protease TW7 encoded the complete
amino acid sequence of the mature protein and 12 amino
acids of a putative "pro" region. The gene encoding the
mature protease TW7 is interrupted by two introns.
15   These introns are of 54 and 84 nucleotides in length.
The exact position of the introns were determined by
comparing the nucleotide sequences of the genomic DNA
with that of the cDNA for protease TW7. The two introns
begin with the nucleotide sequence GT and end with the
20   sequence AG. TAG is used as the termination codon which
is followed by another TAG at an 8 codon interval. The
putative processing of mRNA occurs between sequences
coding for serine (the last amino acid of the possible
pre-pro sequence) and alanine (the first amino acid) of
25   the mature protease TW7.

Example 4

Construction of the cDNA library and isolation of the
30   cDNA gene for protease TW7.

Although the complete gene encoding the mature protease
TW7 was obtained from the genomic library, this gene
cannot be used for the expression of the proteolytic
35   enzyme in microorganisms such as B. subtilis and E. coli
because these organisms lack the enzymes for removing

the intron sequences from the nascent RNA to form the
functional messenger RNA. The isolation of functional
messenger RNA from the T. album Limber fungus was
attempted to obtain a complementary DNA gene, which may

5   be expressed in microorganisms suitable for the
production of protease TW7.

T. album Limber was grown in the skim milk media
described in Example 1 for 9 days, after which mycelia

10   were collected on one layer of miracloth. To about 20 g
of freshly harvested mycelia was added 100 g of sterile
alumina, 45 ml of lysis buffer (4% SDS, 1 mM EDTA, 100 mM
Na acetate, pH 5.0) and 20 ml of phenol:chloroform:isoamyl
alcohol (25:24:1), then mycelia were ground in a sterile

15   mortar with a pestle for 20 min. After addition of 50 ml
of buffer and 50 ml of phenol:chloroform: isoamylalcohol,
the lysate was shaken at room temperature for 30 minutes
before centrifuging at 5000 x g for 15 minutes.
Following phenol:chloroform extraction, 0.1 volume of 3M

20   ammonium acetate and 2.5 volumes of ethanol was added
and the nucleic acids were allowed to precipitate at
-20°C overnight. The polyadenylated RNA species was
isolated using oligo dT-cellulose in accordance with the
procedures described by Maniatis et al. (1982, supra).

25   In order to prevent RNA breakdown during the
manipulation of samples, all buffers were treated with
0.1% diethyl pyrocarbonate and then autoclaved for 60
minutes. The isolated mRNA population was precipitated
with 2.5 volumes of ethanol following the addition 0.1

30   volumes of 3M ammonium acetate.

Northern blot analysis of the mRNA population was
conducted following the glyoxal-DMSO procedure described
by Maniatis et al. (1982). The mRNA species were

35   separated on a 1.1% agarose gel, blotted onto a gene
screen plus membrane (NEF-976, New England Nuclear) and

- 20 -

hybridized with a kinased oligomer probe representing
the amino terminus of protease TW7 and having the
nucleotide sequence:

5' TGGGGCGTCTTCCTGGGTGGC 3'

5   The prehybridization and hybridization were carried out
at 55°C followed by stringent washing at 55°C and
60°C. Upon exposure of the filter to x-ray film, a
single mRNA species of about 2000 nucleotide bases long
was identified in the mRNA population to contain the
10 sequence coding for protease TW7.

Double stranded cDNA was synthesized on polyA+ mRNA
template following the procedures described by Okayama
et al., (1982), Mol. Cell Biol., 2, 161-170. Competent
15 E. coli HB101 cells were transformed with the plasmid
vectors containing the cDNA inserts (Hanahan, 1983, J.
Mol. Biol. 166, 557-580). Transformed colonies on
nitrocellulose filters were replica plated onto a second
set of nitrocellulose filters. The master and the
20 replica filters were incubated at 37°C on L agar plates
containing 50 µg ampicillin/ml, until the colonies grew
up (generally 2-3 hours for master plates and 5-6 hours
for replica). Master filters were stored at 4°C while
the replica filters were incubated overnight on L agar
25 plates containing 100 µg chloramphenicol/ml for plasmid
amplification. Replica filters were processed for DNA
denaturation, renaturation, baking, prehybridization
followed by hybridization with the oligomeric probe as
used for the mRNA detection for Northern blot.
30 Following a second series of screening in which the
isolated single positive colonies were identified,
plasmid DNA was prepared in accordance with the
procedures of Birnboim, (1983, supra). Nucleotide
sequencing of the positive clones was conducted
35 essentially as described for the genomic clone in
accordance with Example 3.

- 21 -

Example 5

Characterization of the cDNA gene for protease TW7.

5   The nucleotide sequence of the cDNA clone has been
       determined from single stranded and double stranded
       DNA. Figure 2 represents the nucleotide sequence of the
       cDNA for protease TW7. The sequence is 1016 bases in
       length and ends with a polyA tail. The restriction map
10   of the clone reveals the presence of an EcoRI site
       proximal to the 5' end of the sequence coding for the
       mature, secreted form of the gene product.

       The nucleotide sequence of the cDNA clone for the
15   protease TW7 gene is identical to the genomic clone,
       except for the lack of introns in the cDNA clone.

Example 6

20   Construction of the cDNA library and isolation of the
       cDNA gene for protease TW3.

       The isolation of functional messenger RNA from the T.
       album fungus was attempted to obtain a complementary DNA
25   gene which could be eventually expressed in industrial
       microorganisms for large scale production of protease
       TW3.

       T. album was grown in a BSA media for 15 days after
30   which mycelia were collected on one layer of
       miracloth. To about 20 g of freshly harvested mycelia
       100 g of sterile alumina and 45 ml of lysis buffer (4%
       SDS, 1 mM EDTA, 100 mM Na acetate, pH 5.0), 20 ml of
       phenol:chloroform:isoamyl alcohol (25:24:1) were added
35   then mycelia were ground in a sterile mortar with a
       pestle for 20 min. Following addition of 50 ml of

- 22 -

buffer, and 50 ml of phenol:chloroform: isoamylalcohol,
the lysate was shaken at room temperature for 30 minutes
before centrifuging at 5000 x g for 15 minutes.
Following extraction with phenol:chloroform, 0.1 volume
5   of 3M ammonium acetate and 2.5 volumes of ethanol were
added and the nucleic acids were allowed to precipitate
at -20°C overnight.

The isolation of polyadenylated RNA species was carried
10   out using oligo dT-cellulose in accordance with the
procedures described by Maniatis et al. (1982, supra).
The isolated mRNA population was precipitated with 2.5
volumes of ethanol following the addition of 0.1 volumes
of 3M ammonium acetate.
15

Northern blot analysis of the mRNA population was
conducted following the glyoxal-DMSO procedure described
by Maniatis et al. (1982). mRNA species were separated
on a 1.1% agarose gel, blotted onto a gene screen plus®
20   membrane (NEN-976, New England Nuclear) and hybridized
with kinased oligomer probes, prepared in Example 2.
The prehybridization and hybridization were conducted at
55°C followed by stringent washing at 55°C and 60°C.
Upon exposure of the filter to x-ray film, a single mRNA
25   of about 2000 nucleotide bases long was identified in
the mRNA population to contain the sequence coding for
protease TW3.

In order to prevent RNA breakdown all buffers used in
30   these examples were treated with 0.1% diethyl
pyrocarbonate and autoclaved for 60 minutes.
Double stranded cDNA was synthesized on polyA+ mRNA
template following the method of Okayama et al. (1982)
<u>Mol. Cell Biol.</u>, <u>2</u>, 161-170. Competent <u>E. coli</u> HB101
35   cells were transformed with the plasmid vectors
containing the cDNA inserts, (Hanahan, 1983, <u>J. Mol.</u>

- 23 -

Biol. 166, 557-580). Transformed colonies on
nitrocellulose filters were replica plated onto a second
set of nitrocellulose filters. The master and the
replica filters were incubated at 37°C on L agar plates
5 containing 50 µg ampicillin/ml, until the colonies grew
up (generally 2-3 hours for master plates and 5-6 hours
for replica). Master filters were stored at 4°C while
the replica filters were incubated overnight on L agar
plates containing 100 µg chloramphenicol/ml for plasmid
10 amplification.

Replica filters were processed for DNA denaturation,
renaturation, baking, prehybridization followed by
hybridization with the oligomeric probe used for the
15 mRNA detection for Northern blot. After a second series
of screening in which the isolated single positive
colonies were identified, plasmid DNA was prepared
(Birnboim, 1983, Methods in Enzymol. 101C, 20-75).
Nucleotide sequencing of the positive clones was carried
20 out as follows: A complete restriction analysis of the
clone was followed by subcloning different fragments
into M13mp18 and mp19 for single strand DNA sequencing.
Double strand sequencing was carried out by the use of
sequence-specific primers. The dideoxy chain termination
25 procedure (Sanger, F. et al., 1977, Proc. Natl. Acad.
Science USA, 74: 5463) was used for DNA sequencing.

Example 7

30 Characterization of the cDNA gene for protease TW3.

A full length cDNA clone was obtained for the protease
TW3. The clone codes for the mature protease as well as
the putative prepro region of the protease. There are
35 four ATG in the open reading frame preceding the
sequence coding for the mature protein. Most probably

- 24 -

the first ATG codes for the initial methionine as it is
followed by an area enriched in hydrophobic amino acids
(Von Heijne, G. 1986, Nucleic Acids Res. 14, 4683-
4690). The putative pro region is comprised of about
5   100 amino acids, a situation very similar to in
subtilisin (Stahl et al., 1984 supra).

The amino acid sequence of the mature protein as
determined from the nucleotide sequence has a large
10   percentage of homology with that of proteinase K. There
is approximately 90% homology between these two
proteases. There are certain positions where the amino
acid resembles that in subtilisins, but not to
proteinase K. For example, at positions 143, a
15   methionine residue occurs in all subtilisins as well as
in protease TW3, while a leucine residue is present at
that position in proteinase K. Similarly at position
219, an alanine residue is present in protease TW3 and
subtilisins, but not in proteinase K. In addition, the
20   amino acid fragment, Ser-Thr-, is absent from proteinase
K while being present in all others in Figure 9 at
position 226 and 227.

Example 8

25

Determination of proteolytic activity.

The proteolytic activity of serine proteases in the
following Examples was determined using azoalbumin as a
30   substrate. Aliquots (20 μl) of enzyme solution or
enzyme buffer (for controls) were mixed with 1 ml of
0.6% azoalbumin in 0.05M Tris-HCl pH 8.2 (unless
otherwise mentioned) and the hydrolytic reaction was
conducted at room temperature. The reaction was
35   terminated after 20 min. upon addition of 10%
trichloracetic acid (400 μl). The hydrolysate was

- 25 -

separated from the precipitated protein by
centrifugation and its optical density at 410 nm, as
compared to the control was measured.

5   Protease activity was also assayed using azocasein as
the substrate. To a final volume of 500 μl, 20 μl of
azocasein (5% solution in 0.2M Tris-HCl, pH 7.5, 1mM
$CaCl_2$), 20 μl of enzyme (1-10 μg) and 460 μl of 50 mM
Tris-HCl, pH 7.5 were added. The samples were incubated
10   at 37°C for 30 min. Following incubation, 500 μl of 10%
TCA was added to the samples and the samples were
incubated on ice for 15 min. After centrifugation for 2
min. 800 μl of supernatant was added to a tube
containing 200 μl of 1.8N NaOH. The optical density of
15   the sample was then measured at 420 nm against the
control.

Example 9

20   Isolation and purification of protease TW7 from the
culture broth of Tritirachium album Limber.

Protease TW7 was separated and purified from the
extracellular media of the T. album culture. A 0.5
25   liter aliquot of the T. album culture broth obtained
after 15 days of fermentation in the BSA-glucose media
described in Example 1 was centrifuged for 10 min.
Proteins in the clear supernatant were precipitated
using ammonium sulfate (180 g) and collected by
30   centrifugation. The precipitate was suspended in 0.02 M
sodium phosphate, pH 6.0 (50 ml) and insoluble material
was removed by centrifugation. The proteins in the
supernatant were reprecipitated with acetone (2.5
volumes). The precipitate was collected by
35   centrifugation and collected on a sintered glass
funnel. The precipitate was then dissolved in water (40

- 26 -

ml) and the resulting solution was dialyzed at 4°C
against 0.02 M sodium phosphate at pH 6.0. The dialyzed
solution was cleared by centrifugation and then passed
through a column (2.5 x 10 cm) of carboxylmethyl-
5   cellulose (CM-52, Whatman) at a rate of 2 ml per minute,
followed by washing the column with 0.02M sodium
phosphate pH 6.0 (30 ml). The flow through and washing
solutions were combined and the proteins were
precipitated by adding 2.5 volumes of acetone. The
10   precipitate was separated by centrifugation, filtered,
and dried under vacuum. The acetone powder (185 mg) was
dissolved in 0.02 M sodium acetate pH 5.0 (2 ml) and
loaded onto a Sephadex G-75 molecular sieve column (2.5
x 90 cm). Fractionation on the molecular sieve column
15   was conducted with 0.02 M sodium acetate pH 5.0 at a
flow rate of 6 ml per hour. Fractions of 2 ml were
collected and monitored for u.v. absorbance at 275 nm.
Of the three major peaks that were eluted from the
column, only one showed proteolytic activity in
20   hydrolyzing the chromogenic protein-substrate
azoalbumin. The enzyme in this peak (between 102-110
ml) was precipitated with acetone (2.5 volumes) and
collected by centrifugation. The precipitate was
dissolved in 2 mM calcium acetate (20 ml) and the
25   solution was dialyzed at 4°C against water, and then
lyophilized. SDS-PAGE of reduced sample demonstrated
protease TW7 as one major band (>95%) with an apparent
molecular weight of 35,000 daltons.

30   Example 10

Amino terminus analysis of protease TW7.

Protease TW7 was purified from the culture supernatant
35   as described in Example 9. The purified protein was
then inactivated with PMSF (phenylmethylsulfonyl-

- 27 -

fluoride) and further purified using HPLC. The amino
terminus was determined by reconstituting the protein in
50% trifluoroacetic acid (TFA) containing 1 mM PMSF.
Automated gas phase Edman degradation was carried out
5   for the amino terminus analysis and the amino terminus
was determined to be Ala-Thr-Gln-Glu-Asp-Ala-Pro-Trp-
Leu-Ala-Arg-Ile-Ser-Ser.

Example 11

10

pH-profile of proteinase TW7.

The pH-profile of proteinase TW7 at 25°C was determined
using azoalbumin as a substrate. Sustrate solutions
15   containing 0.6% azoalbumin in sodium phosphate-sodium
borate buffers covering the pH range between 5.0 and
12.75 were prepared. To the azoalbumin solutions was
added either 10 µl of a 1 mg/ml solution of proteinase
TW7 in water or 10 µl of water (control) and the
20   proteolytic activity of each solution was determined as
described in Example 8. The pH-profile of proteinase
TW7 is represented in Figure 5 wherein the percentage of
maximum proteolytic activity has been plotted versus
pH. The optimum pH range of proteinase TW7 has been
25   determined to be betweena pH range of from 9 to 11.

Example 12

Temperature profile of proteinase TW7.

30

The temperature profile of proteinase TW7 in 0.05 sodium
phosphate at pH 8.5 was determined by measuring the
proteolytic activity over a temperature range of between
20°C and 70°C. Aliquots (1 ml) of 0.6% azoalbumin in
35   0.05M sodium phosphate pH 8.5 were incubated at various
temperatures and after the addition of the enzyme

- 28 -

solution (20 µl of 0.5 mg proteinase TW7 per ml of the same buffer) or 20 µl of buffer (control) the reaction was allowed to proceed for 10 minutes and then terminated upon addition of 10% trichloroacetic acid

5   (400 µl). The temperature profile of proteinase TW7 is represented in Figure 6 wherein the percentage of maximum enzyme activity has been plotted versus temperature. As illustrated in Figure 6, the preferred temperature of proteinase TW7 ranges between 57°C to

10   62°C.

                              Example 13

Determination of protease stability in the presence of

15   detergent.

The stability of protease TW7 was compared in detergent as well as non-detergent solutions with that of subtilisin Carlsberg and to that of proteinase K.

20   Solutions of the enzymes to be evaluated were prepared in appropriate buffers so that the initial proteolytic activities of the various enzyme solutions were similar as measured by the azoalbumin assay described in Example 8. The solutions were incubated at 52°C and aliquots

25   were drawn and measured for residual enzyme activity. The results obtained are represented in Tables 1 and 2. The residual enzyme activity is expressed as a percent of the initial enzyme activity.

30

35

- 29 -

Table 1

Stability of protease TW7 vs. proteinase K and subtilisin
Carlsberg in 0.1M sodium phosphate pH 8.0 with and without SDS.

Without SDS:

Protease	0 Hours	1 Hour	2 Hours
Protease TW7	100%	106%	102%
Proteinase K	100%	67%	43%
Subtilisin Carlsberg	100%	29%	8%

With 0.5% SDS:

Protease	0 Hours	1 Hour	2 Hours	24 Hours
Protease TW7	100%	99%	99%	92%
Proteinase K	100%	32%	9.5%	0%
Subtilisin Carlsberg	100%	5%	0%	0%

Table 2

Stability of protease TW7 vs. Subtilisin Carlsberg in 0.1M sodium
glycinate pH 10.30 containing 0.5% SDS

Protease	t = 0	t = 1 Hr.	t = 18 Hr.	$T_{\frac{1}{2}}$
Protease TW7	100%	105%	80%	5.6 Hr.
Subtilisin Carlsberg	100%	11.3%	0%	0.3 Hr.
Subtilisin BPN	100%	6.6%	0%	0.25 Hr.

- 30 -

Example 14

Stability of protease TW7 in detergent compositions.

5   The stability of protease TW7 and proteinase K were
     evaluated in three enzyme-containing commercial laundry
     detergents, Era Plus® (manufactured by Procter and
     Gamble), Tide® (manufactured by Procter and Gamble) and
     Dynamo® (manufactured by Colgate-Palmolive). The
10   concentrated stock detergents were diluted ten times in
     deionized water. The endogenous protease was
     inactivated by incubating the diluted detergent at 65°C
     for 1 hour prior to addition of the protease to be
     evaluated. The relative activity of the protease to be
15   evaluated in each detergent formulation was adjusted to
     be similar to the enzymatic activity present in the
     original detergent formulation.

     Deactivated detergent containing the proteases to be
20   evaluated were incubated at 52°C along with the
     detergent containing active endogenous enzyme. The
     residual proteolytic activity after different incubation
     periods were quantitated by withdrawing samples and
     assaying in accordance with the procedures as described
25   in Example 8.

     Protease TW7 was found to be stable in all detergent
     formulations tested. For example, in a formulation
     containing deactivated Era Plus®, protease TW7 retains
30   100% of the enzyme activity after 6 hours of incubation
     compared to only 12% of the endogenous enzyme. In a
     formulation containing deactivated Dynamo®, protease TW7
     retained 76% of the enzyme activity after 29 hours of
     incubation.

35

- 31 -

The stability of the protease TW7 was also tested in a formulation containing the laundry detergent Wisk® which does not contain enzymes in its formulation. Following a 1:10 dilution of the stock detergent in deionized

5   water, proteinase K or protease TW7 were added and the resulting composition was incubated at 52°C for various durations. The remaining activity was assayed using the procedures described in Example 8. Protease TW7 was more stable than proteinase K, as illustrated ·in Table

10   4, wherein after 3.5 hours of incubation, proteinase K retained only 20.1% of the original activity while protease TW7 retained approximately 90% of its original enzymatic activity.

15

20

25

30

35

- 32 -

Table 3

Detergent Formulation
(Enzyme)

Time

Era Plus®

	3 h	6 h
(Protease TW7)	95	90
(Proteinase K)	79	64
(Endogenous enzyme)	36	11

Tide®

	1 h	2 h	24 h
(Protease TW7)	107	104	64
(Proteinase K)	58	23	0

Dynamo®

	29 h
(Protease TW7)	76
(Proteinase K)	34.4
(Endogenous enzyme)	22

Wisk®

	1.3 hr	3.5 hr
(Protease TW7)	95.7	89.5
(Proteinase K)	44.3	20.1

Experiments were carried out as described in
Example 14. Percentage of the original activity
remaining after different time points are depicted.

- 33 -

Example 15

Expression of proteinase TW7.

5   An oligonucleotide adaptor was synthesized to insert the
gene coding the mature protease TW7 in the expression
vector pCFM 1156 for expression in E. coli. The adaptor
has the following sequence:
AATTCTAGAAGGAGGAATAACATATGGCCACCCAGGAAGACGCCC

10   GATCTTCCTCCTTATTGTATACCGGTGGGTCCTTCTGCGGGGTAC
The construction of the recombinant plasmid for
expression in E. coli involved the following steps:
(1) Construction of a double-stranded oligonucleotide,
which contains the portion of the expression vector pCFM

15   1156 from the XbaI site up to the NdeI site including
the ATG sequence, followed by the amino terminal
sequence of the mature protease TW7 gene starting from
GCC up to the first NCoI site. Additional nucleotides
were added on the 5' end to obtain an extra EcoRI site

20   that will facilitate the construction.

(2) As illustrated in the Figure 12, the cDNA for
protease TW7 was digested completely with the
restriction enzyme EcoRI and partially with NCoI to

25   obtain the large EcoRI-NcoI fragment, to which kinased
EcoRI-NcoI, adapter was ligated to obtain a cDNA gene
for protease TW7 with the adaptor.

(3) The reconstructed cDNA for protease TW7 was then

30   digested with EcoRI and Cla to obtain the gene for the
protease TW7, which was inserted into M13mp18 which was
already digested with EcoRI and AccI.

(4) The M13mp18 containing the protease TW7 gene was

35   digested with the restriction enzymes XbaI and HindIII
to rescue the protease TW7 gene.

- 34 -

(5) The expression vector pCFM 1156 was also digested
with the restriction enzymes XbaI and HindIII to insert
the protease TW7 gene obtained from M13 mp18 as a XbaI
and HindIII fragment.

5

The resulting recombinant expression plasmid thus
constructed contained a pL promoter, a Shine Dalgarno
sequence, an ATG followed by the nucleotide sequence
coding for the mature protease TW7 protein.

10

Competent $\underline{E.}$ $\underline{coli}$ FMII cells (ln^-, $ptr3^-$) were
transformed with this recombinant plasmid.

Plasmid DNA was isolated from the transformed cells and
15   the positive clones were identified by digesting the
plasmid DNA with XbaI and HindIII. One of the positive
clones was grown overnight to inoculate a medium of Luria
broth containing 20 µg kanamycin/ml. Cells were grown
at 30°C up to an optical density of 0.25 at 600 nm. The
20   temperature of incubation was then shifted to 42°C for 2
hours for the induction of plasmid born gene products.

$\underline{E.}$ $\underline{coli}$ cells were collected by low speed
centrifugation. The pellet was weighed and then
25   suspended in 10 volumes of 50 mM Tris-HCl, pH 7.5. The
cells were lysed by three passes through a French
press. The pellet fraction obtained after another
centrifugation was extracted with 5M urea, 50 mM Tris-
HCl, pH 8.0. The urea soluble proteins were analyzed
30   using various techniques.

The proteins were reduced with β mercaptoethanol and
electrophoresed in a 10% polyacrylamide gel in the
presence of SDS. A unique 35,000 dalton protein was the
35   major band present in the urea soluble fraction derived
from $\underline{E.}$ $\underline{coli}$ harboring the recombinant plasmid but not

- 35 -

the vector alone. It comigrated with the protease TW7
purified from <u>T. album</u>. This protein also reacted
specifically with the antibody raised against the fungal
protease TW7 in a Western blot analysis. This protein

5   was isolated from polyacrylamide gel to identify the
sequences in the amino terminal region, which matched
with that of the fungal protease TW7.

The recombinant protease TW7 is enzymatically inactive

10   when isolated from <u>E. coli</u> prior to refolding. For
reactivation, the protein was suspended in 8M urea, 10
mM DTT, 25 mM Tris-HCl, pH 8.5 in a final concentration
of 1 mg protein per ml and then bound to DE52 resin
preequillibrated with 25 mM Tris-HCl, pH 8.5 at room

15   temperature. Urea was removed by washing the resin with
the same buffer without and then with 4 mM glutathione
(reduced) and 0.4 mM glutathione (oxidized). After
overnight incubation of the resins in the presence of
the above reagents, i.e., 25 mM Tris-HCl, pH 8.5, 4 mM

20   glutathione (reduced) and 0.4 mM glutathione (oxidized),
the protein was eluted from the column with 50 mM Tris-
HCl, pH 7.5 conaining 0.3M NaCl and precipitated with 3
volumes of acetone. This renatured protein had protease
activity against the chromogenic substrate (Example 1)

25   and casein.

Example 16

<u>Isolation and purification of proteinase TW3 from the</u>
30   <u>culture broth of Tritirachium album Limber</u>.

0.5 liter of <u>Tritirachium album</u> Limber culture broth
obtained after 15 days of fermentation in the BSA-
glucose media described in Example 1 was centrifuged at

35   15,000 g for 10 min. Proteins in the clear supernatant
were precipitated with ammonium sulfate (180 g) and

- 36 -

collected by centrifugation. The precipitate was
suspended in 0.02 M sodium phosphate pH 6.0 (50 ml) and
following removal of insoluble material by
centrifugation the proteins in the supernatant were
5   reprecipitated with acetone (2.5 volumes). The
precipitate was collected by centrifugation and
collected on a sintered glass funnel. The precipitate
was dissolved in water (40 ml) and the solution was
dialyzed at 4°C against 0.02 M sodium phosphate at pH
10   6.0. The dialyzed solution was cleared by
centrifugation and passed through a column (2.5 x 10 cm)
of carboxylmethyl-cellulose (CM-52, Whatman) at a rate
of 2 ml per minute, followed by washing the column with
0.02M sodium phosphate pH 6.0 (30 ml). The flow through
15   and washing solutions were combined and the proteins
were precipitated by adding 2.5 volumes of acetone.

Elution from the CM-52 column was accomplished with a
linear gradient of 0 to 0.4M NaCl in sodium phosphate pH
20   6.0. Fractions were assayed spectrophotometrically (at
420 nm) for proteolytic activity at pH 8.2 as described
in Example 8. Peak fractions containing the enzyme
activity were pooled, dialyzed at 4°C against water and
then lyophilized. SDS-PAGE of 2-mercaptoethanol treated
25   sample showed protease TW3 as a major band (>96% based
on Coomassie blue staining) having an apparent molecular
weight of 31,000 daltons.

Example 17

30

Amino terminus analysis of protease TW3.

Protease TW3 was purified from the culture supernatant
as described in Example 16. Purified protein was then
35   inactivated with PMSF (phenylmethylsulfonylfluoride) and
further purified using HPLC. The amino terminus was

- 37 -

determined by reconstituting the protein in 50%
trifluoroacetic acid (TFA) containing 1 mM PMSF.
Automated gas phase Edman degradation was carried out
for the amino terminus analysis. The amino terminus was
5   found to be Ala-Glu-Gln-Arg-Asn-Ala-Pro-Trp-Gly-Leu-Ala-
Arg-Ile-Ser-Ser-Thr.

Example 18

10   <u>pH-profile of proteinase TW3</u>.

The pH-profile of protease TW3 at 25°C was determined
using azocasein as a substrate. Sustrate solutions
(0.6%) M buffers covering the pH range of between 5.0
15   and 10 were made. To these azocasein solutions either
20 µl of an 1 mg/ml solution of protease TW3 in water or
20 µl of water (control) was added and the proteolytic
activity was determined as described in Example 8. The
pH-profile of proteinase TW3 is represented in Figure 10
20   wherein the percentage of maximum activity is plotted
versus pH. The pH optimum of protease TW3 was
determined to be pH 9.

Example 19

25

<u>Temperature profile of proteinase TW3</u>.

The temperature profile of protease TW3 in 0.05 sodium
phosphate at pH 8.5 was determined from the proteolytic
30   activity (azocasein assay) in the temperature range
between 4°C and 65°C. For these measurements, aliquots
(1 ml) of 0.6% azocasein in 0.05M sodium phosphate pH
8.5 were incubated at various temperatures and following
the addition of the enzyme solution (20 µl of 0.5 mg
35   proteinase TW3 per ml of the same buffer) or 20 µl of
buffer alone (control) the reaction was allowed to

- 38 -

proceed for 10 minutes and then terminated upon addition
of 10% trichloroacetic acid (400 μl). The temperature
profile of protease TW3 is represented in Figure 11
wherein the percentage of maximum proteolytic activity

5   has been plotted versus temperature. As illustrated in
Figure 5, the optimum temperature of protease TW3 ranges
between 55°C to 60°C.

                          Example 20

10

Comparative Stability Study of the Protease TW3 and
Subtilisin.

Protease TW3 and a commercial subtilisin (Sigma,
15   protease VII, Cat. No. 5255) were tested for their
stability at 52°C in the absence of substrates. Tests
were carried out at three different pH values (4.0, 8.0
and 10.0).

20   Approximately 0.2 mg of enzyme/ml was incubated at
50°C. The remaining activity after a defined time
interval was quantitated by withdrawing 10 μl of sample
and assaying the proteolytic activity as outlined in
Example 8. Table 4 represents the results obtained.
25   The data represents the retained enzymatic activity as a
percent of the original enzymatic activity. Protease
TW3 demonstrated the best stability. For example, at pH
8.0, after one hour of incubation at 50°C, 90% of the
original activity of protease TW3 was retained, while
30   only 2% of the original activity of subtilisin was
retained. Following 2 hours of incubation, 96% of
protease TW3 activity remained, compared to 0.6% of the
activity of the subtilisin. Protease TW3 was also more
stable than subtilisin at pH 4.0 and 10.0.

35

- 39 -

<u>Table 4</u>

pH 8.0:		Time	
Enzyme	<u>1h</u>	<u>2h</u>	<u>3h</u>
TW3	89.5	96.3	87.4
Subtilisin	2.1	0.6	0.8

pH 10.0:			
Enzyme	<u>1h</u>	<u>2h</u>	<u>3h</u>
TW3	96.8	82.7	74
Subtilisin	30.5	9.8	3.5

pH 4.0:			
Enzyme	<u>1h</u>	<u>2h</u>	<u>3h</u>
TW3	49.3	36	20.5
Subtilisin	0	0	0

- 40 -

Example 21

Stability of Protease TW3 in Commercial Laundry
Detergents.

5

The stability of protease TW3, as well as of subtilisin
were tested in three enzyme-containing commercial
laundry detergents. These were Era Plus® (manufactured
by Procter and Gamble), Tide® (manufactured by Procter
10   and Gamble) and Dynamo® (manufactured by Colgate-
Palmolive). The concentrated stock detergents were
diluted 200 fold in deionized water. The endogenous
protease was inactivated by incubating the diluted
detergent at 65°C for 1 hour, prior to adding either
15   protease TW3, protease TW7, proteinase K or subtilisin
to achieve the enzyme activity present in the original
detergent formulation.

Inactivated detergent with added proteases was incubated
20   at 50°C for various time points along with the detergent
containing active endogenous enzyme. The residual
proteolytic activity remaining after different
incubation periods was quantitated by withdrawing
samples and assaying in accordance with the procedures
25   described in Example 8.

As illustrated in Table 5, protease TW3 is very stable
in all detergent formulations tested. For example, in a
formulation containing Era Plus®, 94% of its activity
30   was retained after 1 hour of incubation while the
formulation containing endogenous enzyme was completely
inactivated. In a formulation containing Dynamo, 80% of
the activity of protease TW3 remained after one hour of
incubation, compared to only a nondetectable level of
35   activity in a formulation containing endogenous
enzyme. In a formulation containing Tide®, 48% of the

- 41 -

original activity of TW3 was retained, while 23% of the
activity in the formulation containing endogenous enzyme
remained after one hour. In all instances subtilisin
was inactivated within one hour.

5

The stability of the protease TW3 was also tested in the
laundry detergent Wisk® which does not contain enzyme in
its formulation. Following a 1:200 dilution of the
stock detergent in deionized water, subtilisin or
10   protease TW3 were added to the diluted detergent and the
samples were incubated at 50°C. The proteolytic
activity was assayed using the procedure outlined in
Example 8. The results obtained are represented in
Table 5.

15

20

25

30

35

- 42 -

Table 5

Detergent Formulation (Enzyme)	Time			
	10'	20'	30'	60'
ERA Plus®				
(TW7)	83.4	94.5	87.7	89
(TW3)	87.7	89.1	92.2	94.8
(Subtilisin)	3.9	1.3	1.5	1.8
(Endogenous enzyme)	50.0	37.7	24.7	0
(Proteinase K)	83.2	84.1	86.7	79.7
Tide®				
(TW7)	103.5	111	99.6	110
(TW3)	78	55.7	68	48
(Subtilisin)	1.8	0.8	0.74	0.5
(Endogenous enzyme)	22.6	22	21.9	23.5
(Proteinase K)	83	75	62	35.4
Dynamo®				
(TW7)	84.9	90	85.9	89.3
(TW3)	83	90.8	82	87.5
(Subtilisin)	30	10.8	3.6	0.5
(Endogenous enzyme)	52	32.5	15.1	0
(Proteinase K)	81.3	83	79.9	87.4
Wisk®				
(TW7)	89.1	87.1	90.8	90
(TW3)	58.3	43.5	28.2	6
(Subtilisin)	21	36	0.7	0.5
(Endogenous enzyme)	0	0	0	0
(Proteinase K)	57.6	39.7	19.7	2.8

- 43 -

Various other examples and modifications of the
foregoing description and examples will be apparent to a
person skilled in the art after reading the present
disclosure without departing from the spirit and scope
5   of the invention, and it is intended that all such
examples or modifications be included within the scope
of the appended claims.

10

15

20

25

30

35

- 44 -

WHAT IS CLAIMED IS:

1. A purified serine protease isolated from culture
    media of <u>Tritirachium album</u> Limber strain ATCC
5   22563.

2. A serine protease according to Claim 1 having an
    amino acid sequence as set forth in Figure 2.

10   3. A serine protease according to Claim 1 having an
    amino acid sequence as set forth in Figure 8.

4. An oligonucleotide probe capable of hybridizing with
    a gene encoding a serine protease isolated from
15   culture media of fungus <u>Tritirachium album</u> Limber
    strain ATCC 22563.

5. An oligonucleotide probe according to Claim 4 having
    the following nucleotide sequence:
20   5' GCT GCT AIT CCG GCA ACG TGA GGA GTC GCC ATG GAC GTT CC 3'

6. A DNA sequence for use in securing expression in a
    procaryotic or eucaryotic host cell a serine
    protease having the structural conformation of a
25   serine protease isolated from culture of media of
    <u>Tritirachium album</u> Limber strain ATCC 22563.

7. A DNA sequence according to Claim 6 having a
    nucleotide sequence set forth in Figure 1.
30
8. A DNA sequence according to Claim 7 having a
    nucleotide sequence set forth in Figure 7.

9. A composition comprising an effective amount of a
35   serine protease isolated from a culture medium of
    fungus <u>Tritirachium album</u> in a detergent composition.

- 45 -

10. A composition according to Claim 9 wherein the serine protease has an amino acid sequence represented in Figure 2.

5   11. A composition according to Claim 9 wherein the serine protease has an amino acid sequence represented in Figure 8.

12. An expression vector capable, in a transformed cell
10   culture, of expressing a DNA sequence according to Claim 6.

13. A cell culture transformed with an expression vehicle according to Claim 12.

15

14. A microorganism according to Claim 13 obtained by transforming an E. coli strain.

15. An E. coli strain capable of producing a serine
20   protease having an amino acid sequence set forth in Figure 2.

16. An E. coli strain capable of producing a serine protease having an amino acid sequence set forth in
25   Figure 8.

17. A purified and isolated serine protease having the structural conformation of a serine protease isolated from a culture of Tritirachium album Limber
30   strain ATCC 22563 and characterized by being the product of procaryotic or eucaryotic expression of an exogenous DNA sequence.

18. A serine protease according to Claim 17 wherein the
35   exogenous DNA sequence is a cDNA sequence.

- 46 -

19. A serine protease according to Claim 17 wherein the
exogenous DNA sequence is a genomic DNA sequence.

20. A serine protease according to Claim 17 wherein the
5   exogenous DNA sequence is a manufactured DNA
sequence.

21. A serine protease according to Claim 17 having a
structural conformation as set forth in Figure 2.
10
22. A serine protease according to Claim 17 having a
structural conformation as set forth in Figure 8.

23. A process for the production of a serine protease
15   having the structural conformation of a serine
protease isolated from a culture of <u>Tritirachium
album</u> Limber strain ATCC 22563, said process
comprising growing under suitable nutrient
conditions procaryotic or eucaryotic host cells
20   transformed or transfected with a DNA vector
including a DNA sequence useful in securing
expression in a procaryotic or eucaryotic host cell
the desired serine protease and isolating desired
serine protease of the expression of DNA sequences
25   in said vector.

30

35

1/18

```
        10         20         30         40         50         60
GAATTCATTG AACAAGACGC CGTTGTTACC ATCTCCCCCA CCCAGCAACA CGCCCCATGG
CTTAAGTAAC TTGTTCTGCG GCAACAATGG TAGAGGGGGT GGGTCGTTCT GCGGGGTACC

        70         80         90        100        110        120
CGTCTGGCCC CCATCTCCAG CCAGGAACCC GCCCCCACCA CTTATACCTA CCATGACTCT
CCAGACCGGG GGTAGAGGTC GGTCCTTGGG CGGGGGTGGT GAATATGGAT GGTACTGAGA

       130        140        150        160        170        180
GCCGGTACAG GCACCTGCCC ATACATCATC GACACGGGCA TCTACACCAA CCACACTCTA
CGGCCATGTC CGTGGACGGG TATGTAGTAG CTGTGCCCGT AGATGTGGTT GGTGTGAGAT

       190        200        210        220        230        240
ACCTTCTCTC CGACCGAATC CGACCAGATC CCAAATGCTA ATAAATGCTA GGACTTTGCC
TGGAAGAGAG GCTGGCTTAG GCTGGTCTAG GGTTTACGAT TATTTACGAT CCTGAAACGG

       250        260        270        280        290        300
GGTCGTGCCA AGTTCCTCAA GAACTTTGCC GGTGACGGTC AAGACACCGA CGGCAACGGT
CCAGCACGGT TCAAGGAGTT CTTGAAACGG CCACTGCCAG TTCTGTGGCT GCCGTTGCCA

       310        320        330        340        350        360
CACGGCACTC ACCTCCCCGG TACCCTGGGC CGAACAACCT ATGGTGTAGC CAAGAAGACA
GTGCCGTGAG TGGAGGGGCC ATGGGACCCG GCTTGTTGGA TACCACATCG GTTCTTCTGT

       370        380        390        400        410        420
TCTCTCTTTG CTGTCAAGGT CCTCGACGCC AACGGTCAGG GCTCCAAGTA CGTTTTGTGT
AGAGAGAAAC GACAGTTCCA GGAGCTGCGG TTGCCAGTCC CGAGGTTCAT GCAAAACACA

       430        440        450        460        470        480
CCTTTCCTCG TGTTCCACCA CCCTATCTTT CTCCTAATCG TCAATTGTAC TAACACATCG
GGAAAGGAGC ACAAGGTGGT GGGATAGAAA GAGGATTAGC AGTTAACATG ATTGTGTAGC

       490        500        510        520        530        540
CCACCCAACA GCTCCGGCCGT CATCGCAGGC ATGGACTTTG TTACCAAAGA CGCCTCGTCC
GGTGGGTTGT CGAGGCCGCA GTAGCGTCCG TACCTGAAAC AATGGTTTCT GCGGAGCAGG

       550        560        570        580        590        600
CAAAACTGCC CCAAGGCCGT CGTAGTGAAC ATGTCGCTCG GTGGTCCCTC CTCCTCACCC
GTTTTGACGG GGTTCCGGCA GCATCACTTG TACAGCGAGC CACCAGGGAG GAGGAGTGGG

       610        620        630        640        650        660
GTCAACCGCG CCGCCGCCGA AATCACCAGC GCAGGCCTCT TCCTCGCTGT CGCAGCCCGC
CAGTTGGCGC GGCGGCGGCT TTAGTGGTCG CGTCCGGAGA AGGAGCGACA GCGTCGGGCG

       670        680        690        700        710        720
AACGAAGCCA CTGACGCCTC CTCGTCGTCC CCTGCGTCCG AAGAAAGCGC CTGCCACTGTC
TTGCTTCGGT GACTGCGGAG GAGCAGCAGG GGACGCAGGC TTCTTTCGCG GACGTGACAG

       730        740        750        760        770        780
GGCCCAACCG ACAAGACCGA CACCCTGGCC GAGTACTCCA ACTTTGGCAG CGTCGTTGAC
CCGGGTTGGC TGTTCTGGCT GTGGGACCGG CTCATGAGGT TGAAACCGTC GCAGCAACTG

       790        800        810        820        830        840
CTCCTTCCTC CCCGTACCGA TATCAAGTCT ACCTGGAACG ACGGCCGCAC CAAGATTATT
GAGGAAGGAG GGGCATGGCT ATAGTTCAGA TGGACCTTGC TGCCGGCGTG GTTCTAATAA

       850        860        870        880        890        900
TCGGGCACGT CCATGGCTAG CCCACATGTT GCTGGACTGC GTCCGTACTT TTTGGGGCCTT
AGCCCGTGCA GGTACCGATC GGGTGTACAA CGACCTGACG CACGGCATGAA AAACCCGGAA

       910        920        930        940        950        960
GGACAAAAGC TTCAGGCTCT TTGCGACTAC ATCGTTGAGA AGGGTCTCAA GGATGTCATT
CCTGTTTTCC AAGTCCGAGA AACGCTGATG TAGCAACTCT TCCCAGAGTT CCTACAGTAA

       970        980        990       1000       1010       1020
CAGAGTCTGC CCACTGATAC TGCCCAATCTT TTGATCAACA ATGGTGAGGG GTCGGCTTAG
GTCTCACACG GGTGACTATG ACGGTTACAA AACTAGTTGT TACCACTCCC CAGCCGAATC

      1030       1040       1050       1060
ATCGGCTTAG ACTTGCATAC ATAGGGGCAC ATCGATGATG
TAGCCGAATC TGAACGTATG TATCGGGCTG TAGCTACTAC
```

FIG. 1

2/18

```
        10         20         30         40         50         60
CAATTCATTC AACAAGACGC CGTTCTTACC ATCTCCGCCA CCCAGGAACA CGCCCCCATGG
CTTAAGTAAC TTGTTCTGCG GCAACAATGG TAGAGGCGGT GGGTCCTTCT GCGGGGGTACC

        70         80         90        100        110        120
GGTCTCGCCC GCATCTCCAG CCACGAACCC GGCGGCACCA CTTATACCTA CGATGACTCT
CCAGAGCGGG CGTAGAGGTC GGTGCTTGGG CCGCCGTGGT GAATATGGAT GCTACTGAGA

       130        140        150        160        170        180
GCCCGTACAG CCACCTCCGC ATACATCATC GACACGGGCA TCTACACCAA CCACACTGAC
CGGGCATGTC CGTGGACGCG TATGTAGTAG CTGTGCCCGT AGATGTGGTT GGTGTGACTG

       190        200        210        220        230        240
TTTGCCGGTC GTGCCAACTT CCTCAACAAC TTTGCCCGGTC ACGGTCAACA CACCGACGCC
AAACGGCCAG CACGGTTCAA GGAGTTCTTG AAACGGGCAC TGCCAGTTCT GTGGCTGCCG

       250        260        270        280        290        300
AACGGTCACG GCACTCACGT CGCCGGTACC GTGGGCGGAA CAACCTATGG TGTAGCCAAG
TTGCCAGTGC CGTGAGTGCA GCGGCCATGG CACCCGCCTT GTTGGATACC ACATCGGTTC

       310        320        330        340        350        360
AAGACATCTC TCTTTGCTGT CAAGGTCCTC GACCCCAACG GTCAGCGCTC CAACTCCGGC
TTCTGTAGAG AGAAACGACA GTTCCAGGAG CTGCGGTTGC CAGTCCCGAG GTTGAGGCCG

       370        380        390        400        410        420
GTCATCGGCAG GCATGGACTT TGTTACCAAA GACGCCTCGT CCCAAAACTG CCCCAAGCGC
CAGTAGCGTC CGTACCTGAA ACAATGGTTT CTGCGGAGCA GGGTTTTGAC GGGGTTCCCG

       430        440        450        460        470        480
GTCGTAGTCA ACATGTCGCT CGGTCGTCCC TCCTCCTCAG CCGTCAACCG CCCCGCCCCC
CAGCATCAGT TGTACAGCGA GCCACCACGC AGGAGGAGTC GGCAGTTGGC GGGGCGGCGG

       490        500        510        520        530        540
GAAATCACCA GCCCAGGCCT CTTCCTCGCT GTCCCAGCCG GCAACCAAGC CACTCACGCC
CTTTACTGGT CGGGTCCGGA GAAGGAGCGA CAGGGTCGGC CGTTGGTTCG GTGACTGCGG

       550        560        570        580        590        600
TCCTCGTCGT CCCCTGCGTC CGAAGAAACC GCCTCCACTG TCGGCGCAAC CGACAAGACC
AGGAGCAGCA GGGGACGCAG GCTTCTTTGG CGGAGGTGAC AGCCGCGTTG GCTGTTCTGG

       610        620        630        640        650        660
GACACGCTGG CCGAGTACTC CAACTTTGGC ACCGTCGTTC ACCTCCTTGC TCCCCGGTACG
CTGTGCGACC GGCTCATGAG GTTGAAACCG TGGCAGCAAC TGGAGGAACG AGGGGCCATGC

       670        680        690        700        710        720
GATATCAACT CTACCTCGAA CGACGGCCGC ACCAAGATTA TTTCGGGCAC GTCCATGGCT
CTATAGTTGA GATGGAGCTT GCTGCCGGCG TGGTTCTAAT AAAGCCCGTG CAGGTACCGA

       730        740        750        760        770        780
AGCCCACATG TTCCTGGACT CGGTGCGTAC TTTTTCGGCC TTGGACAAAA GGTTCAGCGT
TCGGGTGTAC AACGACCTGA GCCACGCATG AAAAAGCCGG AACCTGTTTT CCAAGTCCCA

       790        800        810        820        830        840
CTTTGCGACT ACATCGTTGA GAAGGGTCTC AAGGATCTCA TTCAGAGTGT GCCCAGTGAT
GAAACGCTGA TGTACCAACT CTTCCCAGAG TTCCTACAGT AAGTCTCACA CGGGTCACTA

       850        860        870        880        890        900
ACTGCCAATG TTTTGATCAA CAATGGTGAG GCCTCGGCTT AGATCGGCTT AGAGTTGCAT
TGACGGTTAC AAAACTAGTT GTTACCACTC CCGAGCCGAA TCTAGCCGAA TCTCAACGTA

       910        920        930        940        950        960
ACATAGCCCC ACATCCATCA TGGCATGTTG GGCGAATTAC TGTATATATT GCACAGTAGA
TGTATCGGGC TGTAGGTAGT ACCCTACAAC CCGCTTAATG ACATATATAA CGTGTCATCT

       970        980        990       1000       1010       1020
CATACACAGT CGTTTTGATA ACGGCCGTTG CATTCAATTC ATCTTCTACT TGATTTAAAA
GTATGTGTCA GCAAAACTAT TGCCGGCAAC GTAAGTTAAG TAGAAGATGA ACTAAATTTT
```

AAAA
TTTT

FIG. 2

```
-12     -10                                                    1
glu phe ile glu gln asp ala val val thr ile ser ala thr gln
CAA TTC ATT GAA CAA GAC GCC GTT GTT ACC ATC TCC GCC ACC CAG

                         10
glu asp ala pro trp gly leu ala arg ile ser ser gln glu pro
GAA GAC GCC CCA TGG GGT CTC GCC CGC ATC TCC AGC CAG GAA CCC

         20                                    30
gly gly thr thr tyr thr tyr asp asp ser ala gly thr gly thr
GGC GGC ACC ACT TAT ACC TAC GAT GAC TCT GCC GGT ACA GGC ACC

                               40
cys ala tyr ile ile asp thr gly ile tyr thr asn his thr asp
TGC GCA TAC ATC ATC GAC ACG GGC ATC TAC ACC AAC CAC ACT GAC

         50                                    60
phe gly gly arg ala lys phe leu lys asn phe ala gly asp gly
TTT GGC GGT CGT GCC AAG TTC CTC AAG AAC TTT GCC GGT GAC GGT

                               70
gln asp thr asp gly asn gly his gly thr his val ala gly thr
CAA GAC ACC GAC GGC AAC GGT CAC GGC ACT CAC GTC GCC GGT ACC

         80                                    90
val gly gly thr thr tyr gly val ala lys lys thr ser leu phe
GTC GGC GGA ACA ACC TAT GGT GTA GCC AAG AAG ACA TCT CTC TTT

                              100
ala val lys val leu asp ala asn gly gln gly ser asn ser gly
GCT GTC AAG GTC CTC GAC GCC AAC GGT CAG GGC TCC AAC TCC GGC

         110                                   120
val ile ala gly met asp phe val thr lys asp ala ser ser gln
GTC ATC GCA GGC ATG GAC TTT GTT ACC AAA GAC GCC TCG TCC CAA

                          . 130
asn cys pro lys gly val val val asn met ser leu gly gly pro
AAC TGC CCC AAG GGC GTC GTA GTC AAC ATC TCG CTC GGT GGT CCC

         140                                   150
ser ser ser ala val asn arg ala ala ala glu ile thr ser ala
TCC TCC TCA GCC GTC AAC CGC GCC GCC GCC GAA ATC ACC AGC GCA

                              160
gly leu phe leu ala val ala ala gly asn glu ala thr asp ala
GGC CTC TTC CTC GCT GTC GCA GCC GGC AAC GAA GCC ACT GAC GCC

         170                                   180
ser ser ser ser pro ala ser glu glu ser ala cys thr val gly
TCC TCG TCG TCC CCT GCG TCC GAA GAA AGC GCC TGC ACT GTC GGC

                              190
ala thr asp lys thr asp thr leu ala glu tyr ser asn phe gly
GCA ACC GAC AAG ACC GAC ACG CTG GCC GAG TAC TCC AAC TTT GGC

         200                                   210
ser val val asp leu leu ala pro gly thr asp ile lys ser thr
AGC GTC GTT GAC CTC CTT GCT CCC GGT ACC GAT ATC AAG TCT ACC

                              220
trp asn asp gly arg thr lys ile ile ser gly thr ser met ala
TGG AAC GAC GGC CGC ACC AAG ATT ATT TCG GGC ACG TCC ATC GCT

         230                                   240
ser pro his val ala gly leu gly ala tyr phe leu gly leu gly
AGC CCA CAT GTT GCT GGA CTC GCT GCC TAC TTT TTC GCC CTT GGA

                              250
gln lys val gln gly leu cys asp tyr met val glu lys gly leu
CAA AAG GTT CAG GGT CTT TGC GAC TAC ATG GTT GAG AAG GGT CTC

         260                                   270
lys asp val ile gln ser val pro ser asp thr ala asn val leu
AAG GAT GTC ATT CAG AGT GTC CCC AGT GAT ACT GCC AAT GTT TTG

                        280 281
ile asn asn gly glu gly ser ala AM
ATC AAC AAT GGT GAG GCC TCG GCT TAG ATCGCCTTAGAGTTGCATACATAG

CCCGACATCCATGATGGGATGTTGGGCCGAATTAGTGTATATATTGCACAGTAGACATACA

GAGTGGTTTTGATAACGGCCCCTTGGATTGAATTGATGTTGTACTTGATTTAAAAAAAAA
```

FIG. 3

4/18

```
                                     10              20   23
    TW7      A T Q E D A P W G L A R I S S Q E P G G T T Y

    PK       A A Q T N A P W G L A R I S S T S P G T S T Y

    NO       A Q S - V P Y G V S Q I K - - - - - A P A L

    CA       A Q T - V P Y G I P L I K - - - - - A D K V

    DY       A Q T - V P Y G I P L I K - - - - - A D K V

    TH     Y T P N D P Y F S S - R Q Y G P Q K I Q - - - - - A P Q A


             24        30            40            50
    TW7      T Y D D S A G T G T C A Y I I D T G I Y T N H T D F - - - -

    PK       Y Y D E S A G Q G S C V Y V I D T G I E A S H P E F - - - -

    NO       H S Q G Y T G S N V K V A V I D S G I D S S H P D L - - K V

    CA       Q A Q G F K G A N V K V A V L D T G I Q A S H P D L - - N V

    DY       Q A Q G Y K G A N V K V G I I D T G I A A S H T D L    K V

    TH       W - D I A E G S G A K I A I V D T G V Q S N H P D L A G K V


             54        60            70            80   83
    TW7      G G R A K F L K N F A G D G Q D T D G N G H G T H V A G T V

    PK       E G R A Q M V K T Y Y Y S S R -   D G N G H G T H C A G T V

    NO       A G G A S M V P S E T P N F Q - - D D N S H G T H V A G T V

    CA       V G G A S F V A G E A - Y N T - - D G N G H G T H V A G T V

    DY       V G G A S F V S G E S - Y N T - - D G N G H G T H V A G T V

    TH       V G G W D F V D N D S T P - Q - - N G N G H G T H C A G I A


             84        90            100           110  113
    TW7      G G T           T Y G V A K K T S L F A V K V L D A N G Q

    PK       G S G R - - - - - T Y G V A K K T Q L F - V K V L D D N G S

    NO       A A L - N N S I G V L G V A P S S A L Y A V K V L G D A G S

    CA       A A L - D N T T G V L G V A P S V S L Y A V K V L N S S G S

    DY       A A L - D N T T G V L G V A P N V S L Y A I K V L N S S G S

    TH       A A V T N N S T G I A G T A P K A S I L A V R V L D N S G S
```

FIG. 4-A

```
             114        120               130              140   143
    TW7  G S N S G V I A G M D F V T K D A S S Q N C P K G V V V N M
    PK   G Q Y S T I I A G M D F V A S D K N N R N C P K G V V A S L
    NO   G Q Y S W I I N G I E W - A I A - N N M D - - - - - V I N M
    CA   G S Y S G I V S G I E W - A T T - N G M D - - - - - V I N M
    DY   G T Y S A I V S G I E W - A T Q - N G L D         V I N M
    TH   G T W T A V A N G I T Y - A A D - Q G A U         V I S L


             144        150               160              170   174
    TW7  S L G G P S S S A V N R A A A - E I T S A G L F L A V A A G N
    PK   S L G G G Y S S S V N S A A A - R L Q S S G V M V A V A A G N
    NO   S L G G P S Q S A A L K A A V D K A V A S G V V V V A A A G N
    CA   S L G G A S G S T A M K Q A F D N A Y A R G V V V V A A A G N
    DY   S L G G P S G S T A L K Q A V D K A Y A S G I V V V A A A G N
    TH   S L G G I V G N S G L Q Q A V N Y A W N K G S V V V A A A G N


             175        180               190              200   204
    TW7  E A T D A S S S S     - P A S E E S A C T V G A T D K T D T
    PK   N N A D A R N Y S - - - P A S E P S V C T V G A S D R Y D R
    NO   E G S T G S S S T V G Y P G K Y P S V I A V G A V D S S N Q
    CA   S G N S G S T N T I G Y P A K Y D S V I A V G A V D S N S N
    DY   S G S S G S Q N T I G Y P A K Y D S V I A V G A V D S N K N
    TH   A G N T A P N - - - - Y P A Y Y S N A I A V A S T D Q N D N


             205        210               220              230   234
    TW7  L A E Y S N F G S V V D L L A P G T D I K S T W N D G R T K
    PK   R S S F S N Y G S V L D I F G P G T D S L - - W I G G S T R
    NO   R A S F S S R G P E L D V M A P G V S I Q S T L P G N K Y G
    CA   R A S F S S V G A E L E V M A P G A G V Y S T Y P T N T Y A
    DY   R A S F S S V G A E L E V M A P G V S V Y S T Y P S N T Y T
    TH   K S S F S T Y G S V V D V A A P G S W I Y S T Y P T S T Y A
```

FIG. 4-B

```
            235        240            250                  260   267
    TW7     I I S G T S M A S P H V A G L G A Y F L G L G Q K V Q G L

    PK      S I S G T S M A T P H V A G L A A Y L M T L G K T T A A S A

    NO      A Y N G T S M A S P H V A G A A A L I L S K H P N W T N T Q

    CA      T L N G T S M A S P H V A G A A A L I L S K H P N L S A S Q

    DY      S L N G T S M A S P H V A G A A A L I L S K Y P T L S A A Q

    TH      S L S G T S M A T P H V A G V A G L L A S Q G R S - - A S N
```

```
            265        270            280                  290        295
    TW7     C D - Y M V E K G L K D V I Q S V P S D T A N V L I N N G E G S A

    PK      C R - Y I A D T A N K G D L S N I P F G T V N L A Y N N Y Q A

    NO      F R S S L Q N T T T K L G D S F Y Y G K G L I N V Q A A A Q

    CA      V R N R L S S T A T Y L G S S F Y Y G K G L I N V E A A A Q

    DY      V R N R L S S T A T N L G D S F Y Y G K G L I N V E A A A Q

    TH      I R A A I E N T A D K I S T G T Y W A K G R V N A Y K A V Q Y
```
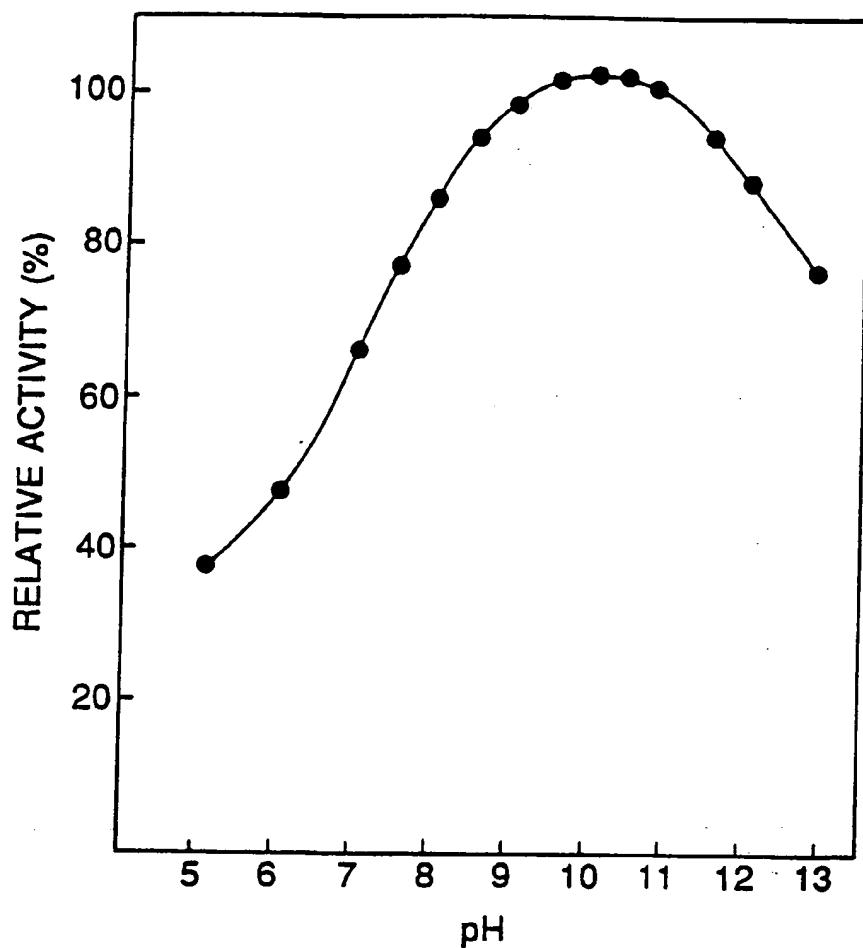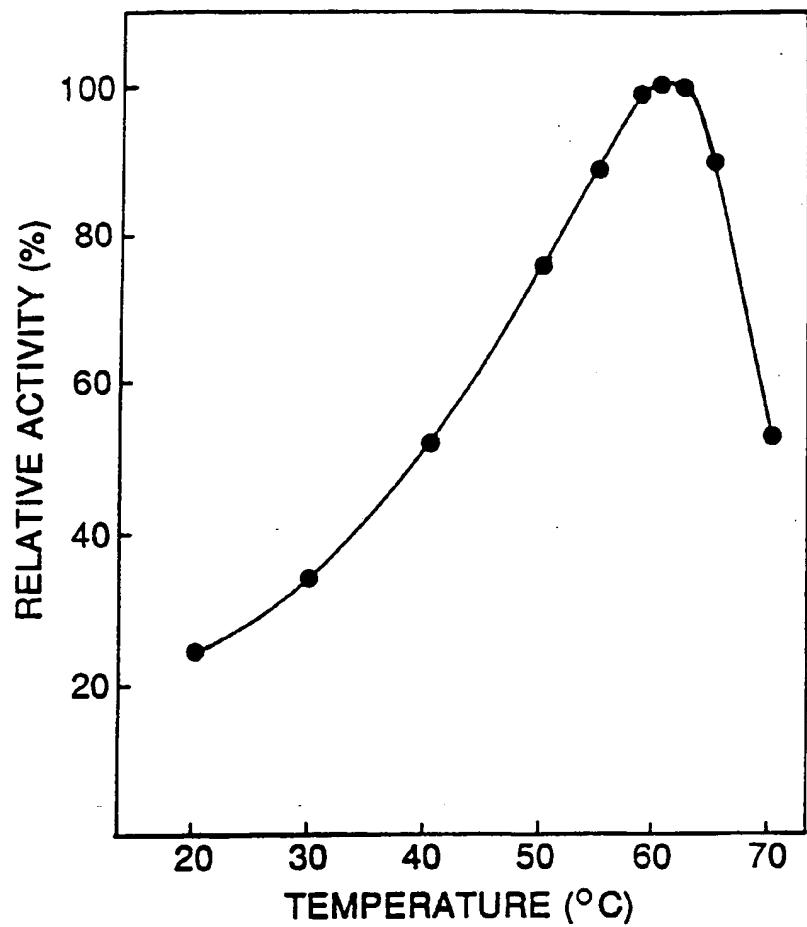
FIG. 4-C

FIG. 5

FIG. 6

9/18

```
        10         20         30         40         50         60
GCGGTTCATC ATCAACACCC ATCGCAGCAA TACAAAAGCG TCTTCTCCAG CTCAACAACA
CCCCAAGTAG TAGTTGTCGG TAGCGTCGTT ATGTTTTCGC AGAAGAGGTC GAGTTGTTGT

        70         80         90        100        110        120
CCTCTTGAAT AAGCCACGCT TTTTCATTCA CCGACGGTCA AAATGCGTCT TTCCATTCTT
GGAGAACTTA TTCGGTGCGA AAAAGTAAGT GGCTGCCAGT TTTACGCAGA AAGGTAAGAA

       130        140        150        160        170        180
CTGGGTCTTC TTCCCCTCGC TCCTCGGCCT CCCGCCGTCG ACGCTGTTGA GCAGCGCTCC
GACCCAGAAG AAGGGGAGCG AGGAGCCGGA GGGCGGCAGC TGCGACAACT CGTCGCGAGG

       190        200        210        220        230        240
GAGCCCGCTC CTCTTATTGA GGGCCAGCGC CACATCATTG CCGACAACTA CATTCTCAAG
CTCGGGCGAG GAGAATAACT CCCGGTCCCG CTCTACTAAC GGCTGTTCAT GTAACAGTTC

       250        260        270        280        290        300
CTCAAGGAGG GTAGCGGCTCT TGCTTCTCTC GATGCTGCCA TGGAGAAGCT TTCTGGCAAG
GAGTTCCTCC CATCGCGAGA ACGAAGAGAG CTACGACGGT ACCTCTTCGA AAGACCGTTC

       310        320        330        340        350        360
GCCGACCACG TCTACAAGAA CATCTTCAAG GGCTTTGCTG CCTCTCTTGA CGAGAAGATG
CGGCTGGTGC AGATGTTCTT GTAGAAGTTC CCGAAACGAC GGAGAGAACT GCTCTTCTAC

       370        380        390        400        410        420
GTTGAGGTCC TCCGCGCCCA CCCTGATGTC GAGTACATTG AGCAGGATCC TATCGTCAAC
CAACTCCAGG AGGCGCGGGT GGGACTACAG CTCATGTAAC TCGTCCTAGG ATAGCAGTTG

       430        440        450        460        470        480
ATCAACGCTG AGCACCGCAA CGCTCCCTGG GGTCTTGCTC GCATCTCCAG CACCAGCCCC
TAGTTGCGAC TCGTGGCGTT GCGAGGGACC CCAGAACGAG CGTAGAGGTC GTGGTCGGGG

       490        500        510        520        530        540
GGTACCTCCA CGTACAGATA CGACGACTCT GCCGCGCCAGG GTACTTGCGT CTACGTCATC
CCATGGAGGT GCATGTCTAT GCTGCTGAGA CGGCGCGGTCC CATGAACGCA GATGCAGTAG

       550        560        570        580        590        600
GACACCGCTG TCGAGGCATC TCACCCCGAG TTTGAGGGCC GCGCCCAGAT GGTCAAGACG
CTGTGGCGAC AGCTCCGTAG AGTGGGGCTC AAACTCCCGG CGCGGGTCTA CCAGTTCTGC
```

FIG. 7-A

```
        610        620        630        640        650        660
   TACTACGCCT CCAGCCGCCA TGGCAACGGC CACGGCACTC ACTGCCCCCG TACCATTGCC
   ATGATGCCGA GGTCGGCGCT ACCGTTGCCG GTGCCGTGAG TGACGCGGCC ATGGTAACCG

        670        680        690        700        710        720
   TCCAGGACCT ACGGTCTCGC CAAGAAGACC CAGATCTTTG GTGTCAAGCT CCTCAACGAC
   AGGTCCTGGA TGCCAGAGCG GTTCTTCTGG GTCTAGAAAC CACAGTTCCA GGAGTTGCTG

        730        740        750        760        770        780
   CAAGGCTCTG GCCAGTACTC CACCATCATC TCTGGTATCG ACTTTGTCGC CAACGACTAC
   GTTCCGAGAC CGGTCATGAG GTGGTAGTAG AGACCATAGC TGAAACAGCG GTTGCTGATG

        790        800        810        820        830        840
   CGCCAACCGCA ACTGCCCCAA CGGTCTCGTT GCCTCCATGT CCATTGGTCG TGGCTACTCC
   GCGTTGGCGT TGACGGGGTT GCCAGAGCAA CGGAGGTACA GGTAACCAGC ACCGATGAGG

        850        860        870        880        890        900
   TCTTCCGTGA ACAGCGCCGC TGCCAACCTC CAGCCAATCTC GTCTCATGGT CGCCGTCGCT
   AGAAGGCACT TGTCGCGGCG ACGGTTGGAG GTCGTTAGAC CACAGTACCA GCGGCAGCGA

        910        920        930        940        950        960
   GCTGGCAACA ACAACGCTGA CGCTCGCAAC TACTCCCCTG CTTCTGAGTC CTCCATCTGC
   CGACCGTTGT TGTTGCGACT GCGAGCGTTG ATGAGGGGAC GAAGACTCAG GAGGTAGACG

        970        980        990       1000       1010       1020
   ACTGTTGGTG CCACTGACCG CTACGACCGA CGATCCAGCT TCTCCAACTA CGGCAGCGTT
   TGACAACCAC GGTGACTGGC GATGCTGGCT GCTAGGTCGA AGAGGTTGAT GCCGTCGCAA

       1030       1040       1050       1060       1070       1080
   TTCGACATCT TTGCCCCCGC TACCGACATT CTCTCCACCT GGATCGGCGG CAGCACCAGA
   AAGCTGTAGA AACGGGGGCG ATGGCTGTAA GAGAGGTGGA CCTAGCCGCC GTCGTGGTCT

       1090       1100       1110       1120       1130       1140
   TCCATCTCTG GTACCTCCAT GGCTACTCCC CACGTTGCTG GTCTCGCTGC CTACCTTATG
   AGGTAGAGAC CATGGAGGTA CCGATGAGGG GTGCAACGAC CAGAGCGACG GATGGAATAC

       1150       1160       1170       1180       1190       1200
   ACTCTCGGAC GCGCCACCGC CAGCAACGCT TGCCGATACA TTGCCCAGAC TGCCAACCAC
   TGAGAGCCTG CGCGGTGGCG GTCGTTGCGA ACGGCTATGT AACGGGTCTG ACGGTTGGTC

       1210       1220       1230       1240       1250       1260
   GGCGATCTGA GCAACATTTC CTTCGGCACT GTCAACCTCC TTGCCTACAA CAACTACCAG
   CCGCTAGACT CGTTGTAAAG GAAGCCGTGA CAGTTGGACG AACGGATGTT GTTGATGGTC

       1270       1280       1290       1300       1310       1320
   CGCTAACTCC TTCAGTCAGC TCTAAAAGTT GGAACATATG AAACGAGATT TGAATGCATC
   GCGATTGAGG AAGTCAGTCG AGATTTTCAA CCTTGTATAC TTTGCTCTAA ACTTACGTAG

       1330       1340       1350       1360       1370       1380
   TGTACATAGA TGAGCACATT CATATCGGTC TTACGACCAT ATTTAGTCAT ATTTAAAAAT
   ACATGTATCT ACTCGTGTAA GTATAGCCAG AATGCTGGTA TAAATCACTA TAAATTTTTA
```

```
   AAA
   TTT
```

FIG. 7-B

GGCGTTCATCATCAACAGCCATCGCAGCAATACAAAAGCGTCTTCTCCAGCTCAACAAC

                                                        -108
                                                met arg leu ser
ACCTCTTGAATAAGCCACGCTTTTTCATTCACCGACGGTCAAA ATG CGT CTT TCC

            -100   -90
ile leu leu gly leu leu pro leu ala pro arg pro pro ala val
ATT CTT CTC GGT CTT CTT CCC CTC GCT CCT CCG CCT CCC GCC GTC

                                    -80
asp ala val glu gln arg ser glu pro ala pro leu ile glu ala
GAC GCT GTT GAG CAG CGC TCC GAG CCC GCT CCT CTT ATT GAG GCC

            -70   -60
gln gly glu met ile ala asp lys tyr ile val lys leu lys glu
CAG GGC GAG ATG ATT GCC GAC AAC TAC ATT GTC AAG CTC AAC GAG

                                        -50
gly ser ala leu ala ser leu asp ala ala met glu lys leu ser
GGT AGC GCT CTT GCT TCT CTC GAT GCT GCC ATG GAG AAG CTT TCT

            -40   -30
gly lys ala asp his val tyr lys asn ile phe lys gly phe ala
GGC AAC GCC GAC CAC GTC TAC AAG AAC ATC TTC AAG GGC TTT GCT

                                        -20
ala ser leu asp glu lys met val glu val leu arg ala his pro
GCC TCT CTT GAC GAG AAG ATG GTT GAG GTC CTC CGC GCC CAC CCT

            -10   1
asp val glu tyr ile glu gln asp ala ile val asn ile asn ala
GAT GTC GAG TAC ATT GAG CAG GAT GCT ATC GTC AAC ATC AAC GCT

                                    10
glu gln arg asn ala pro trp gly leu ala arg ile ser ser thr
GAG CAG CGC AAC GCT CCC TGG GGT CTT GCT CGC ATC TCC AGC ACC

            20   30
ser pro gly thr ser thr tyr arg tyr asp asp ser ala gly gln
AGC CCC GGT ACC TCC ACC TAC AGA TAC GAC GAC TCT GCC GGC CAG

                                    40
gly thr cys val tyr val ile asp thr gly val glu ala ser his
GGT ACT TGC GTC TAC GTC ATC GAC ACC GGT GTC GAG GCA TCT CAC

            50   60
pro glu phe glu gly arg ala gln met val lys thr tyr tyr ala
CCC GAG TTT GAG GGC CGC GCC CAG ATG GTC AAG ACC TAC TAC GCC

                                    70
ser ser arg asp gly asn gly his gly thr his cys ala gly thr
TCC AGC CGC GAT GGC AAC GGC CAC GGC ACT CAC TGC GCC GGT ACC

                                                FIG. 8-A

12/18

```
                    80                                                    90
ile gly ser arg thr tyr gly val ala lys lys thr gln ile phe
ATT GGC TCC AGG ACC TAC GGT GTC GCC AAG AAG ACC CAG ATC TTT


gly val lys val leu asn asp gln gly ser gly gln tyr ser thr
                                   100
GGT GTC AAG GTC CTC AAC GAC CAA GGC TCT GGC CAG TAC TCC ACC


                    110                                         120
ile ile ser gly met asp phe val ala asn asp tyr arg asn arg
ATC ATC TCT GGT ATG GAC TTT GTC GCC AAC GAC TAC CGC AAC CGC


                                             130
asn cys pro asn gly val val ala ser met ser ile gly gly gly
AAC TGC CCC AAC GGT GTC GTT GCC TCC ATG TCC ATT GGT GGT GGC


                    140                                         150
tyr ser ser ser val asn ser ala ala ala asn leu gln gln ser
TAC TCC TCT TCC GTG AAC AGC GCC GCT GCC AAC CTC CAG CAA TCT


gly val met val ala val ala ala gly asn asn asn ala asp ala
                                   160
GGT GTC ATG GTC GCC CTC GCT GCT GGC AAC AAC AAC GCT GAC GCT


                    170                                         180
arg asn tyr ser pro ala ser glu ser ser ile cys thr val gly
CGC AAC TAC TCC CCT GCT TCT GAG TCC TCC ATC TGC ACT GTT GGT


                                             190
ala thr asp arg tyr asp arg arg ser ser phe ser asn tyr gly
GCC ACT GAC CGC TAC GAC CGA CGA TCC AGC TTC TCC AAC TAC GGC


                    200                                         210
ser val leu asp ile phe ala pro gly thr asp ile leu ser thr
AGC GTT TTG GAC ATC TTT GCC CCC GGT ACC GAC ATT CTC TCC ACC


                                             220
trp ile gly gly ser thr arg ser ile ser gly thr ser met ala
TGG ATC GGC GGC AGC ACC AGA TCC ATC TCT GGT ACC TCC ATG GCT


                    230                                         240
thr pro his val ala gly leu ala ala tyr leu met thr leu gly
ACT CCC CAC GTT GCT GGT CTC GCT GCC TAC CTT ATG ACT CTC GGA


                                             250
arg ala thr ala ser asn ala cys arg tyr ile ala gln thr ala
CGC GCC ACC GCC AGC AAC GCT TGC CGA TAC ATT GCC CAG ACT GCC


                    260                                         270
asn gln gly asp leu ser asn ile ser phe gly thr val asn leu
AAC CAG GGC GAT CTC AGC AAC ATT TCC TTC GGC ACT GTC AAC CTG


                                             279
leu ala tyr asn asn tyr gln gly OC
CTT GCC TAC AAC AAC TAC CAG GGC TAA GTCCTTCAGTCACCTCTAAAAGTT

GGAAGATATGAAACGAGATTTGAATGCATCTGTACATAGATGAGCACATTCATATCGGTC

TTACCACCATATTTAGTCATATTTAAAAATAAA                    FIG. 8-B
```

```
                              10                    20   23
    TW3        A E Q R N A P W G L A R I S S T S P G T S T Y

    PK         A A Q T N A P W G L A R I S S T S P G T S T Y

    NO           A Q S - V P Y G V S Q I K - - - - - A P A L

    CA           A Q T - V P Y G I P L I K - - - - - A D K V

    DY           A Q T - V P Y G I P L I K - - - - - A D K V

    TH     Y T P N D P Y F S S - R Q Y G P Q K I Q - - - - - A P Q A
```

```
         24        30                40                  50
    TW3   R Y D D S A G Q G T C V Y V I D T G V E A S H P E F - - - -

    PK     Y Y D E S A G Q G S C V Y V I D T G I E A S H P E F - - - -

    NO     H S Q G Y T G S N V K V A V I D S G I D S S H P D L - - K V

    CA     Q A Q G F K G A N V K V A V L D T G I Q A S H P D L - - N V

    DY     Q A Q G Y K G A N V K V G I I D T G I A A S H T D L     K V

    TH     W - D I A E G S G A K I A I V D T G V Q S N H P D L A G K V
```

```
         54        60                70              80   83
    TW3   E G R A Q M V K T Y Y A S S R - - D G N G H G T H C A G T I

    PK     E G R A Q M V K T Y Y Y S S R -   D G N G H G T H C A G T V

    NO     A G G A S M V P S E T P N F Q - - D D N S H G T H V A G T V

    CA     V G G A S F V A G E A - Y N T - - D G N G H G T H V A G T V

    DY     V G G A S F V S G E S - Y N T - - D G N G H G T H V A G T V

    TH     V G G W D F V D N D S T P - Q - - N G N G H G T H C A G I A
```

```
         84        90              100             110   113
    TW3   G S - R           T Y G V A K K T Q I F G V K V L N D Q G S

    PK     G S G R - - - - - T Y G V A K K T Q L F - V K V L D D N G S

    NO     A A L - N N S I G V L G V A P S S A L Y A V K V L G D A G S

    CA     A A L - D N T T G V L G V A P S V S L Y A V K V L N S S G S

    DY     A A L - D N T T G V L G V A P N V S L Y A I K V L N S S G S

    TH     A A V T N N S T G I A G T A P K A S I L A V R V L D N S G S
```

FIG. 9-A

14/18

```
        114        120                 130              140    143
TW3     G Q Y S T I I S G M D F V A N D Y R N R N C P N G V V A S M

PK      G Q Y S T I I A G M D F V A S D K N N R N C P K G V V A S L

NO      G Q Y S W I I N G I E W - A I A - N N M D - - - - - V I N M

CA      G S Y S G I V S G I E W - A T T - N G M D - - - - - V I N M

DY      G T Y S A I V S G I E W - A T Q - N G L D         V I N M

TH      G T W T A V A N G I T Y - A A D - Q G A U         V I S L


        144        150                 160              170    174
TW3     S I G G G Y S S S V N S A A A - N L Q Q S G V M V A V A A G N

PK      S L G G G Y S S S V N S A A A - R L Q S S G V M V A V A A G N

NO      S L G G P S Q S A A L K A A V D K A V A S G V V V V A A A G N

CA      S L G G A S G S T A M K Q A F D N A Y A R G V V V V A A A G N

DY      S L G G P S G S T A L K Q A V D K A Y A S G I V V V A A A G N

TH      S L G G I V G N S G L Q Q A V N Y A W N K G S V V V A A A G N


        175        180                 190                200    204
TW3     N N A D A R N Y S - - - P A S E S S I C T V G A T D R Y D R

PK      N N A D A R N Y S - - - P A S E P S V C T V G A S D R Y D R

NO      E G S T G S S S T V G Y P G K Y P S V I A V G A V D S S N Q

CA      S G N S G S T N T I G Y P A K Y D S V I A V G A V D S N S N

DY      S G S S G S Q N T I G Y P A K Y D S V I A V G A V D S N K N

TH      A G N T A P N - - - - Y P A Y Y S N A I A V A S T D Q N D N


        205        210                 220              230    234
TW3     R S S F S N Y G S V L D I F A P G T D I L S T W I G G S T R

PK      R S S F S N Y G S V L D I F G P G T D S L - - W I G G S T R

NO      R A S F S S R G P E L D V M A P G V S I Q S T L P G N K Y G

CA      R A S F S S V G A E L E V M A P G A G V Y S T Y P T N T Y A

DY      R A S F S S V G A E L E V M A P G V S V Y S T Y P S N T Y T

TH      K S S F S T Y G S V V D V A A P G S W I Y S T Y P T S T Y A
```

FIG. 9-B

```
         235      240           250              260   267
TW3   S I S G T S M A T P H V A G L A A Y L M T L G R A T A S N A

PK    S I S G T S M A T P H V A G L A A Y L M T L G K T T A A S A

NO    A Y N G T S M A S P H V A G A A A L I L S K H P N W T N T Q

CA    T L N G T S M A S P H V A G A A A L I L S K H P N L S A S Q

DY    S L N G T S M A S P H V A G A A A L I L S K Y P T L S A A Q

TH    S L S G T S M A T P H V A G V A G L L A S Q G R S - - A S N
```

```
         265      270              280              290        295
TW3   C R - Y I A Q T A N Q G D L S N I S F G T V N L A Y N N Y Q G

PK    C R - Y I A D T A N K G D L S N I P F G T V N L A Y N N Y Q A

NO    F R S S L Q N T T T K L G D S F Y Y G K G L I N V Q A A A Q

CA    V R N R L S S T A T Y L G S S F Y Y G K G L I N V E A A A Q

DY    V R N R L S S T A T N L G D S F Y Y G K G L I N V E A A A Q

TH    I R A A I E N T A D K I S T G T Y W A K G R V N A Y K A V Q Y
```
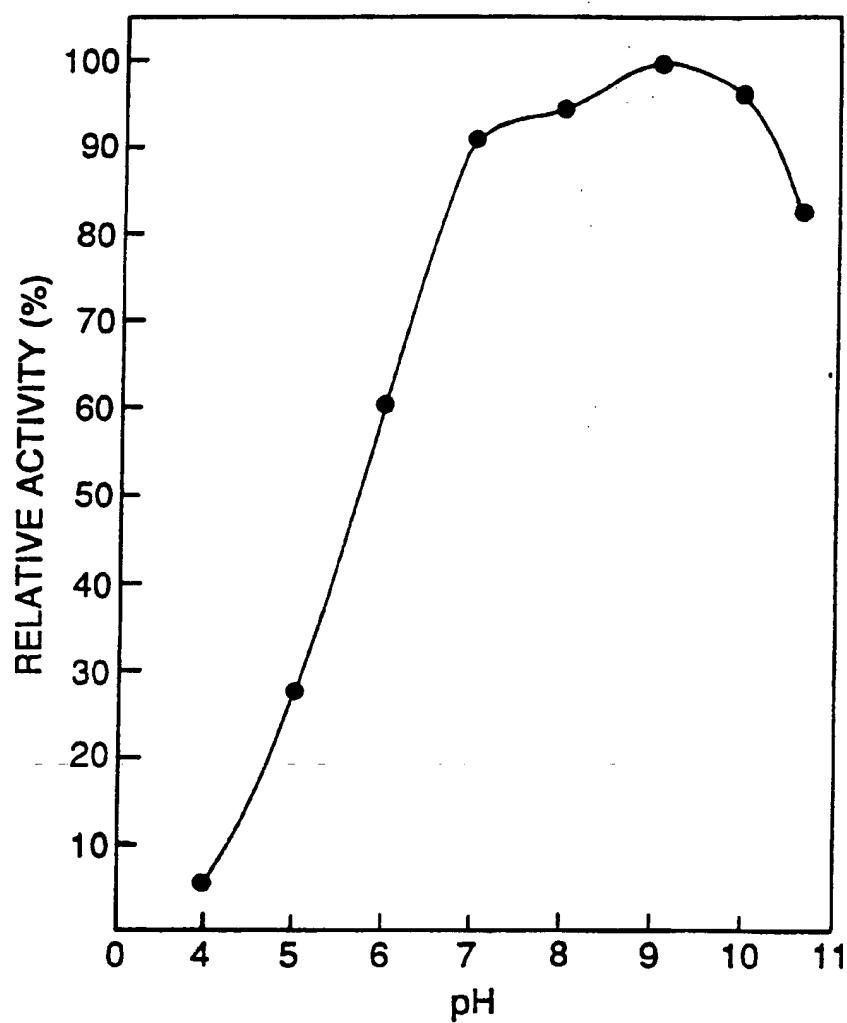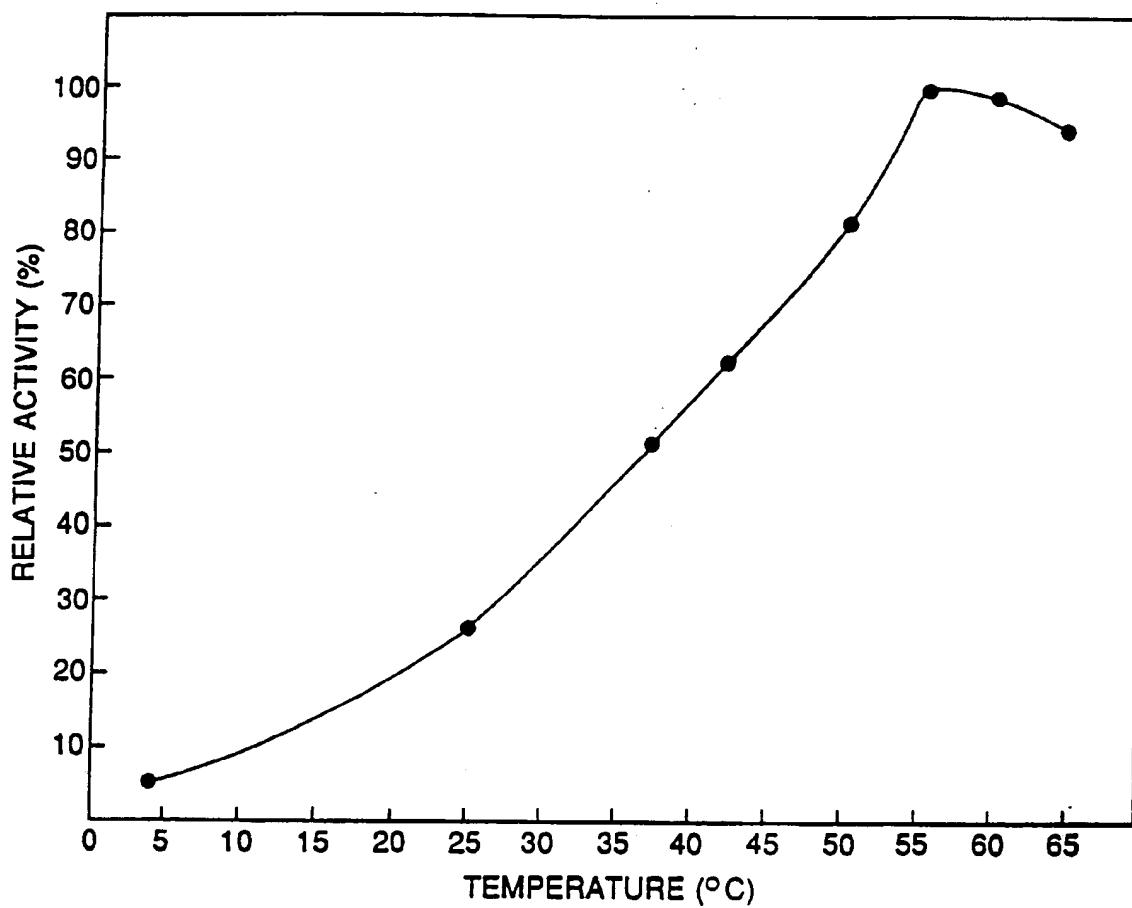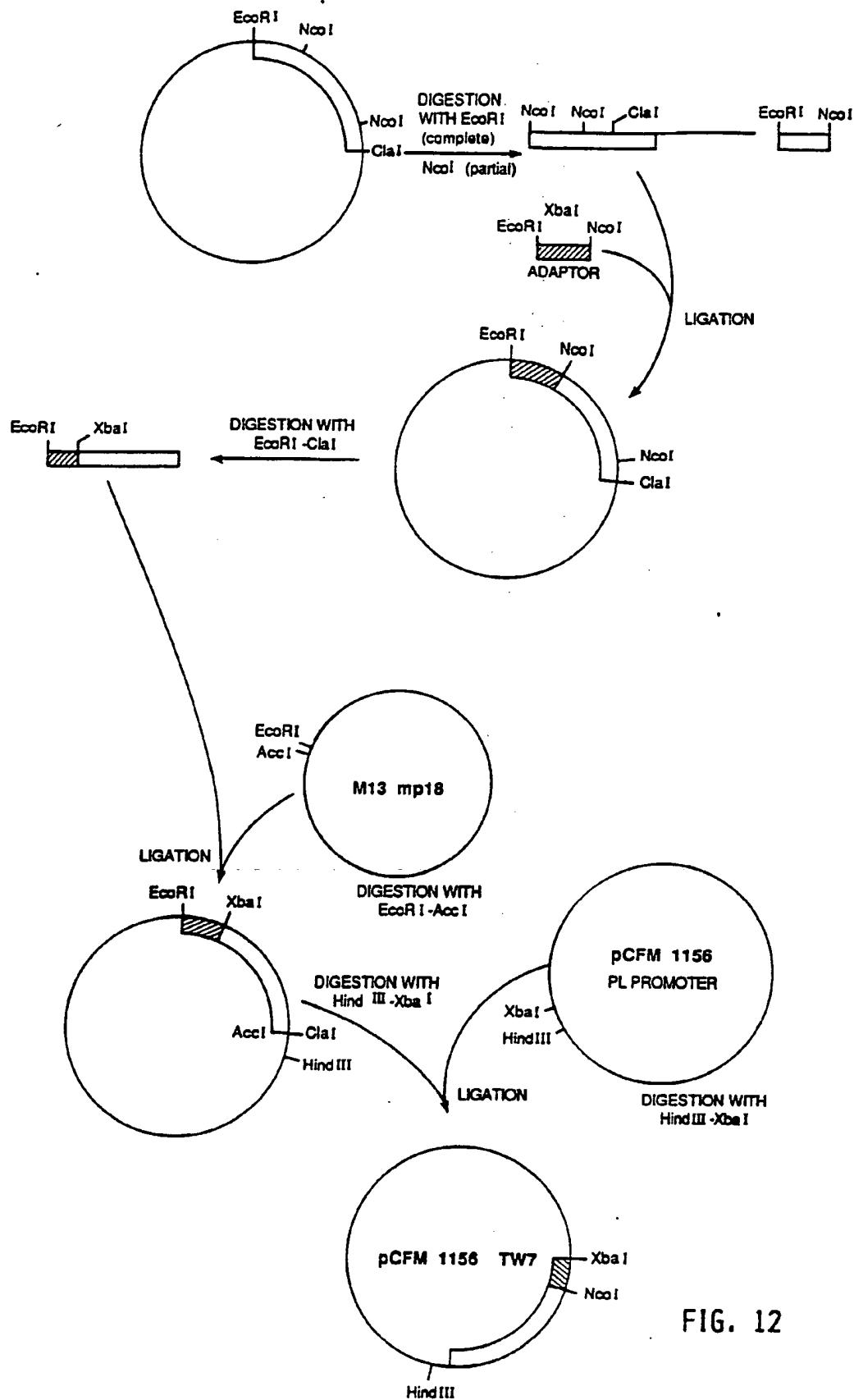
FIG. 9-C

FIG. 10

FIG. 11

18/18

FIG. 12

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US88/01040

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): C12P 21/00; C12N 15/00; C12N 9/58; See Attachment
US CL : 435/68,172.3,223,240,253,254,320; See Attachment

II. FIELDS SEARCHED

Minimum Documentation Searched [7]	
Classification System	Classification Symbols
U.S.	435/68,91,172.3,223,240,253,320; 252/174.12 536/27 935/14,28,29,68,73

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched [8]

CHEMICAL ABSTRACTS DATA BASE (CAS) 1967-1988; BIOLOGICAL ABSTRACTS
DATA BASE (BIOSIS) 1967-1988. KEYWORDS: PLASMID, CLONING,
SERINE PROTEASE, TRITIRACHIUM ALBUM.

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [*]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y	FEBS LETTERS, Volume 199, No. 2, issued April 1986, (Amsterdam, NL), (JANY ET AL), "Amino Acid Sequence of Proteinase K from the Mold Tritirachium album Limber," See pages 139-44.	1-8 and 12-23
Y	Chemical Abstracts, Volume 103, No. 1, issued 8 July 1985 (Columbus, Ohio, USA), (JANY ET AL) "Proteinase K from Tritirachium album Limber. I. Molecular Mass and Sequence Around the Active Site Serine," see page 272, column 1, the abstract No. 2890j, Biol. Chem. Hoppe-Seyler. 1985, 366(5),485-92 (Eng.).	1-23

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
08 June 1988	1 2 JUL 1988
International Searching Authority	Signature of Authorized Officer
ISA/US	Thomas Mays

Attachment To Form PCT/ISA/210, Part I

Part I: CLASSIFICATION OF SUBJECT MATTER:

IPC(4):   C12N 5/00; C12N 1/20; C07H 15/12;
           C11D 7/42

US CL :   536/27;

III. DOCUMENTS CONSIDERED TO BE RELEVANT   (CONTINUED FROM THE SECOND SHEET)

Category*	Citation of Document, ¹⁶ with indication, where appropriate, of the relevant passages ¹⁷	Relevant to Claim No ¹⁸
Y	EMBO JOURNAL, Volume 3, No. 6, issued June 1984 (Oxford, England, UK) (PAHLER ET AL) "Three-Dimensional Structure of Fungal Proteinase K Reveals Similarity to Bacterial Subtilisin", see pages 1311-14.	1-8 and 12-23
Y	Chemical Abstracts, Volume 83, issued 1975 (Columbus, Ohio, USA), "Stimulation of Proteinase K Action by Denaturing Agents. Application to the Isolation of Nucleic Acids and the Degradation of Masked Proteins," see page 236, column 1, the abstract No. 110435s, Eur. J. Biochem. 1975, 56(1), 103-8 (Eng.).	1-23
Y	US, A, 3,790,482 (JONES) 5 February 1974. See the entire document and in particular columns 6 and 7.	9-11
A	US, A, 3,623,957 (FELDMAN) 30 November 1971. See entire document.	1-23

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V.☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.☐ Claim numbers   . because they relate to subject matter [12] not required to be searched by this Authority, namely:

2.☐ Claim numbers   . because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

3.☐ Claim numbers_____. because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

Group I. Claims 1-8 and 12-23 drawn to composition of serine protease, DNA expression vector and cloning host and process of producing serine protease classified in Classes 435 and 536 subclasses 68, 172.3,223,240,253,254 and 320; 27 respectively.
See Attachment

1.☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.   Telephone Practice

2.☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3.☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4.☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (supplemental sheet (2) (Rev. 11-87)

PCT/US88/01040

Attachment to Form PCT/ISA/210, Part VI

Part VI: OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING:

Group II. Claims 9-11 drawn to composition of detergent
and serine protease classified in Class 252 subclass
174.12.